

UNIVERSIDAD DE CUENCA
FACULTAD DE INGENIERÍA
ESCUELA DE INFORMÁTICA



TÍTULO:

**“MODELAMIENTO DEL PERFIL DE UN USUARIO USANDO
TECNOLOGIAS SEMANTICAS”**

Tesis previa a la
obtención del título de
Ingeniero de Sistemas

ELABORADO POR:

Sebastián Rodrigo Román Jarrín

Juan José Sáenz Peñafiel

DIRECTOR:

Ing. Jorge Mauricio Espinoza Mejía, MSc. PhD

mauricio.espinoza@ucuenca.edu.ec

Cuenca – Ecuador

2014

RESUMEN

Este trabajo presenta los procesos y actividades realizadas para el modelamiento de un perfil de usuario de Televisión Digital utilizando tecnologías semánticas y recursos externos que permitan enriquecer esta información. La estructura de este documento está dividida en 6 capítulos.

El primer capítulo contiene la definición del proyecto, alcance y objetivos de la tesis, para posteriormente en el capítulo 2 detallar el contexto tecnológico en el que se basa este proyecto. Dentro de este capítulo se presenta el marco teórico y conceptual de los temas a ser tratados dentro del desarrollo de la tesis que son Perfiles de Usuario y su modelamiento, Redes Sociales, Ontologías y Metodologías para su Creación. Luego, en el capítulo 3 se detalla paso a paso la metodología de creación de ontologías, denominada NeOn que fue seleccionada dentro del contexto tecnológico. Posteriormente, en el capítulo 4 se presenta los métodos usados para la recolección de datos de perfil de usuario desde Redes Sociales, formulario de recolección de datos, y eventos de televisión digital que son considerados fuentes externas.

Como siguiente capítulo, el capítulo 5 detalla todo el procedimiento seguido para la construcción del prototipo de recolección de datos de perfiles de usuario mediante tecnologías semánticas, de este capítulo se trata los temas de la configuración de repositorio, la definición de arquitectura de la aplicación y el proceso de análisis, diseño, implementación y pruebas del prototipo definitivo, en el cual se tendrá los resultados obtenidos en el proyecto, tanto en análisis previo como en construcción de prototipo. Para finalizar con el capítulo 6 de conclusiones y recomendaciones. Adicionalmente en este capítulo 6 se plantean trabajos futuros dentro de la misma línea de orientación de perfiles de usuario y tecnologías semánticas.

Palabras claves: TDT (Televisión Digital Terrestre), RDF, SPARQL, Web Semántica, Redes Sociales, Perfil de Usuario, Acces Token, Repositorio Semántico, Oauth.

ABSTRACT

The processes and activities carried out for the modeling of a digital television user profile which apply semantic technologies and external resources, are presented in this research to enrich this information. The structure of this research is divided in six chapters.

The definition, scope, and goals of the project are contained in the first chapter. Subsequently, the technological context, on which the project is based, is detailed in the second chapter. This chapter involves the theoretical and conceptual framework of the topics to be dealt in the development of the thesis. Among these topics are User profiles and their modeling, Social networks, Ontologies, and Methodologies for their creation. Then, in the third chapter, the Ontology creation methodology called NeOn, which was selected within the technological context, is detailed step-by-step. Later, the method used to the collection of user profiles data from Social Networks, the data collection form, and the digital television events as external sources are specified in the chapter four.

The next chapter, the chapter five, details the entire procedure followed to the prototype construction of the collection of user profiles data through semantic technologies. The configuration repository topics, the definition of the application architecture and the analytical process, design, implementation, and testing of a definitive prototype, on which will be presented the results achieved in the project, are addressed both in previous analysis as in prototype construction. Finally, conclusions and recommendations are presented in the chapter six. Furthermore, in this chapter, future works within the same guideline of user profiles and semantic technologies are outlined.

Keywords: TDT (Terrestrial Digital Television), RDF, SPARQL, Semantic Web, Social Network, User Profile, Acces Token, Semantic Repository, Oauth.

INDICE

AGRADECIMIENTOS.....	16
DEDICATORIAS.....	17
CAPÍTULO 1	19
INTRODUCCIÓN.....	19
1.1. IDENTIFICACIÓN DEL PROBLEMA.....	20
1.2. JUSTIFICACIÓN.....	21
1.3. ALCANCE	22
1.4. OBJETIVO GENERAL.....	23
1.5. OBJETIVOS ESPECÍFICOS	23
1.6. ESTRUCTURA DE LA TESIS	23
CAPÍTULO 2	26
CONTEXTO TECNOLÓGICO	26
2.1. PERFIL DE USUARIO.....	26
2.1.1. Definición	26
2.1.2. Información de usuario.....	27
2.1.2.1. Extracción implícita	27
2.1.2.2. Extracción explícita	28
2.1.2.3. Extracción híbrida	28
2.1.3. Modelamiento de perfiles de usuario	29
2.1.3.1. Modelado del perfil con preferencias explícitas	29
2.1.3.2. Modelado del perfil basado en contenidos.....	29
2.1.3.3. Modelamiento del perfil basado en la interacción entre pares.....	30
2.1.3.4. Modelamiento del perfil basado en el conocimiento.....	31
2.1.3.5. Modelamiento Híbrido de perfil	31
2.1.4. Creación de un modelo perfil de usuario	31
2.2. REDES SOCIALES	33
2.2.1. Definición	33
2.2.2. Selección de redes sociales	34
2.3. ONTOLOGÍAS	37



UNIVERSIDAD DE CUENCA

2.3.1. Ontologías y su representación del conocimiento	38
2.3.2. Elementos de una ontología	39
2.3.3. Metodologías de creación de ontologías	40
2.3.3.1. CYC	41
2.3.3.2. USCHOLD Y KING	41
2.3.3.3. GRÜNINGER Y FOX	42
2.3.3.4. KACTUS	42
2.3.3.5. METHONTOLOGY	43
2.3.3.6. NeOn	43
2.3.3.7. Análisis comparativo de las metodologías	45
CAPÍTULO 3	48
CREACIÓN DE LA ONTOLOGÍA DE PERFIL DE USUARIO	48
3.1. INTRODUCCIÓN	48
3.2. ESPECIFICACIÓN DE REQUERIMIENTOS DE LA ONTOLOGÍA	49
3.3. MODELAMIENTO DEL PERFIL DEL USUARIO	56
CAPÍTULO 4	67
RECOLECCIÓN DEL PERFIL DE USUARIO DESDE FUENTES EXTERNAS	67
4.1. METODOLOGIA	67
4.2. TÉCNICAS PARA RECOLECCIÓN DE DATOS	68
4.2.1. Información recolectada con técnicas explícitas.	68
4.2.2. Información recolectada con técnicas implícitas.	69
4.3. RECOLECCIÓN DE DATOS EXPLÍCITOS	69
4.3.1. Formulario Web, Datos personales	72
4.3.2. Formulario Web, Preferencias	73
4.3.3. Esquema general de los formularios Web	75
4.4. RECOLECCIÓN DE DATOS IMPLÍCITOS	75
4.4.1. Redes sociales	76
4.4.1.1. Especificaciones técnicas de las Redes sociales.	77
4.4.1.2. Protocolos de comunicación	78
4.4.1.3. Procesos de autenticación	80
4.4.1.4. Formatos de respuesta	83
4.4.2. MÓDULO DE REDES SOCIALES MRS 1.0	85

4.4.2.1. Librería HybridAuth	85
4.4.2.2. Configuración librería HybridAuth.....	85
4.4.2.3. API RestFull	91
4.4.3. Eventos Televisión Digital	96
CAPÍTULO 5	99
PROTOTIPO PERFIL DE USUARIO	99
5.1. ANÁLISIS.....	99
5.1.4.1. Caso de uso CU-01 Autenticación de Usuario.....	100
5.1.4.2. Caso de uso CU-02 Registro de Usuario mediante formulario.....	101
5.1.4.3. Caso de uso CU-03 Registro de Usuario mediante redes sociales	102
5.1.4.4. Caso de uso CU-04 Registro de Usuario mediante formulario y redes sociales	103
5.1.4.5. Caso de uso CU-05 Modificación de preferencias del usuario.....	104
5.2. DISEÑO	105
5.2.1. Arquitectura interna.....	106
5.3. REPOSITORIO	107
5.3.1. SESAME	107
5.3.2. APACHE JENA.....	108
5.3.3. VIRTUOSO.....	108
5.3.3.1. Ejecución local del servidor VIRTUOSO	109
5.3.3.2. Entorno web servidor VIRTUOSO.....	111
5.3.3.3. Carga de modelo ontológico al repositorio Virtuoso.....	111
5.3.4 Acceso remoto al repositorio VIRTUOSO	112
5.4. IMPLEMENTACIÓN.....	113
5.4.1. Plataforma de desarrollo web.....	113
5.4.2. Implementación de formularios web.....	115
5.4.3. Implementación de Redes sociales.....	117
5.4.4. Implementación de la Aplicación Web	118
5.4.5. Implementación de la recolección de datos desde la TV.....	120
5.5. PRUEBAS Y RESULTADOS.....	122
5.5.1. Fundamentos de la prueba.....	122
5.5.1.1. Pruebas del prototipo	122



UNIVERSIDAD DE CUENCA

5.5.1.2. Pruebas de Consistencia en la Ontología.....	132
CAPÍTULO 6	143
CONCLUSIONES Y TRABAJOS FUTUROS.....	143
6.1. CONCLUSIONES.....	143
6.2. TRABAJOS FUTUROS	144
REFERENCIAS	146

INDICE DE FIGURAS

FIGURA 1 ESQUEMA GENERAL DEL PROYECTO “APLICACIÓN DE TECNOLOGÍAS SEMÁNTICAS PARA DISMINUIR LA SOBRECARGA DE INFORMACIÓN EN USUARIOS DE TV DIGITAL”	19
FIGURA 2 MODELO DE UN PERFIL DE USUARIO EN ICQ [50]	32
FIGURA 3 EJEMPLO DE ONTOLOGÍA	40
FIGURA 4 TAREAS PARA REQUISITOS ONTOLOGÍA ESPECIFICACIÓN	51
FIGURA 5 METAMODELO DE LA ONTOLOGÍA DEL PERFIL DE USUARIO	65
FIGURA 6 TÉCNICA DE RECOLECCIÓN DE INFORMACIÓN HÍBRIDA	67
FIGURA 7 DISEÑO DEL FORMULARIO WEB PARA LOS DATOS PERSONALES DEL USUARIO	74
FIGURA 8 DISEÑO DEL FORMULARIO WEB PARA LAS PREFERENCIAS DEL USUARIO	74
FIGURA 9 PORCENTAJE DE POBLACIÓN CON CELULAR Y REDES SOCIALES SEGÚN INEC	77
FIGURA 10 FLUJO DEL PROTOCOLO OAUTH	79
FIGURA 11 FLUJO DEL PROTOCOLO OPENID RP (CLIENTE), OP (PROVEEDOR OPENID)	80
FIGURA 12 FLUJO DE SOLICITUD DE ACCESO DE FACEBOOK	81
FIGURA 13 FLUJO DE SOLICITUD DE ACCESO DE TWITTER	82
FIGURA 14 FLUJO DE SOLICITUD DE ACCESO DE GOOGLE+	82
FIGURA 15 FORMATO DE RESPUESTA DE SOLICITUD DE FACEBOOK	83
FIGURA 16 FORMATO DE RESPUESTA DE SOLICITUD DE TWITTER	84
FIGURA 17 FORMATO DE RESPUESTA DE SOLICITUD DE GOOGLE+	84
FIGURA 18 PANEL DE ADMINISTRACIÓN DE FACEBOOK	86
FIGURA 19 PANEL DE ADMINISTRACIÓN DE GOOGLE+	86
FIGURA 20 PANEL DE ADMINISTRACIÓN DE TWITTER	86
FIGURA 21 ESTRUCTURA DE ARCHIVOS LIBRERÍA HYBRIDAUTH	87
FIGURA 22 CONFIGURACIÓN DEL ARRAY DE REDES SOCIALES DE LA LIBRERÍA HYBRIDAUTH	87
FIGURA 23 ESTRUCTURA DE CLASES LIBRERÍA HYBRIDAUTH	89
FIGURA 24 FORMATO DE RESPUESTA JSON DEL PERFIL DE USUARIO EN FACEBOOK	93
FIGURA 25 FORMATO DE RESPUESTA JSON DE LOS “LIKES” DE USUARIO EN FACEBOOK	93
FIGURA 26 FORMATO DE RESPUESTA JSON DE LOS CONTACTOS DE USUARIO EN TWITTER (A QUIEN SIGUE)	94
FIGURA 27 FORMATO DE RESPUESTA JSON DE LAS ACTIVIDADES DEL USUARIO EN GOOGLE+	94
FIGURA 28 ESTRUCTURA TABLA SESIONES HYBRIDAUTH	95
FIGURA 29 RESPUESTA DEL MÓDULO A UNA SOLICITUD REST POR PRIMERA VEZ	95
FIGURA 30 ACCESO A LOS DATOS DE LA TELEVISIÓN DIGITAL	97
FIGURA 31 DISEÑO DEL PROTOTIPO DE MODELAMIENTO DE USUARIO PARA LA TELEVISIÓN DIGITAL	106
FIGURA 32 ARQUITECTURA DE VIRTUOSO. FUENTE: [53]	108
FIGURA 33 INTERFAZ DE EJECUCIÓN DEL REPOSITORIO VIRTUOSO	110



UNIVERSIDAD DE CUENCA

FIGURA 34 ENTORNO WEB DEL REPOSITORIO VIRTUOSO	110
FIGURA 35 INTERFAZ WEB “LINKED DATA” DEL REPOSITORIO VIRTUOSO.....	111
FIGURA 36 INTERFAZ PARA CARGAR UNA ONTOLOGÍA EN EL REPOSITORIO VIRTUOSO.....	112
FIGURA 37 ONTOLOGÍAS CARGADAS EN EL REPOSITORIO VIRTUOSO	112
FIGURA 38 ACCESO AL REPOSITORIO VIRTUOSO MEDIANTE EL MÓDULO MIT	113
FIGURA 39 FORMULARIOS WEB INTEGRADOS EN LA APLICACIÓN WEB DEL PROTOTIPO	116
FIGURA 40 INTERFAZ CRUD PREFERENCIAS DE USUARIO EN EL PROTOTIPO	116
FIGURA 41 INTERFAZ FORMULARIO WEB PARA DATOS PERSONALES DE USUARIO EN EL PROTOTIPO	117
FIGURA 42 INTERFAZ DE REGISTRO DE DATOS CON REDES SOCIALES	117
FIGURA 43 INTERFAZ DE SOLICITUD DE PERMISOS DE FACEBOOK Y GOOGLE+.	118
FIGURA 44 MÓDULO MRS DEFINIDO EN LA ESTRUCTURA DEL PROTOTIPO	118
FIGURA 45 INTERFAZ DE INICIO DE SESIÓN EN EL PROTOTIPO.....	119
FIGURA 46 INTERFAZ DE REGISTRO DE USUARIO EN EL PROTOTIPO	120
FIGURA 47 INTERFAZ DE PERFIL DE USUARIO EN EL PROTOTIPO	121
FIGURA 48 LOGIN CORRECTO, SALUDO INICIAL DEL PERFIL DE USUARIO EN EL PROTOTIPO	122
FIGURA 49 LOGIN INCORRECTO, MENSAJE DE ERROR AL USUARIO	123
FIGURA 50 MENSAJES DE ERROR EN LA CREACIÓN DEL USUARIO	124
FIGURA 51 INTERFAZ DE REGISTRO DE DATOS PERSONALES CON FORMULARIO WEB.	124
FIGURA 52 MENSAJE DE SOLICITUD PARA LLENAR CAMPOS OBLIGATORIOS	125
FIGURA 53 MENSAJE DE REGISTRO CORRECTO DEL USUARIO.....	125
FIGURA 54 INTERFAZ DE AUTENTIFICACIÓN CON LA RED SOCIAL FACEBOOK.....	126
FIGURA 55 INTERFAZ DE AUTENTIFICACIÓN CON LAS REDES SOCIALES TWITTER Y GOOGLE+.	127
FIGURA 56 MENSAJE DE REGISTRO CORRECTO EN EL MÓDULO MRS.....	127
FIGURA 57 INTERFAZ DE SOLICITUD DE PERMISOS DE FACEBOOK Y GOOGLE+.	127
FIGURA 58 MENSAJE DE REGISTRO INCORRECTO EN EL MÓDULO MRS.	128
FIGURA 59 INTERFAZ PARA EL CRUD DE PREFERENCIAS.	130
FIGURA 60 INTERFAZ PARA AGREGAR NUEVAS PREFERENCIAS	131
FIGURA 61 INTERFAZ PARA EDITAR PREFERENCIAS.....	131
FIGURA 62 MENSAJE PARA ELIMINAR PREFERENCIAS.....	132
FIGURA 63 INTERFAZ PARA REALIZAR CONSULTAS SPARQL EN EL PROTOTIPO.	133
FIGURA 64 PRUEBA PC1. RESPUESTA CORRECTA DEL NOMBRE DEL TELEVIDENTE.	134
FIGURA 65 PRUEBA PC2. RESPUESTA CORRECTA DE LA RESIDENCIA DEL TELEVIDENTE.....	135
FIGURA 66 PRUEBA PC3. RESPUESTA CORRECTA DE LA EDAD DEL TELEVIDENTE.	136
FIGURA 67 PRUEBA PC4. RESPUESTA CORRECTA SOBRE LOS TÓPICOS DE INTERÉS DEL TELEVIDENTE.	137
FIGURA 68 PRUEBA PC5. RESPUESTA CORRECTA DE PROGRAMAS QUE MIRA A LAS 11:00H EL TELEVIDENTE.....	138
FIGURA 69 PRUEBA PC6. RESPUESTA CORRECTA DE PROGRAMAS QUE MIRA EL DÍA DOMINGO “2014-08-06”EL TELEVIDENTE.....	139



UNIVERSIDAD DE CUENCA

FIGURA 70 PRUEBA PC7. RESPUESTA CORRECTA DE LOS PERSONAJES FAVORITOS DEL TELEVIDENTE	140
FIGURA 71 PRUEBA PC8. RESPUESTA CORRECTA DE LA PROFESIÓN DEL TELEVIDENTE.....	141

INDICE DE TABLAS

TABLA 1 TABLA COMPARATIVA DE LAS REDES SOCIALES MÁS IMPORTANTES EN EL MEDIO.....	36
TABLA 2 ANÁLISIS COMPARATIVO DE METODOLOGÍAS.....	46
TABLA 3 PLANTILLA DE ESPECIFICACIÓN DE REQUERIMIENTOS DE LA ONTOLOGÍA .	50
TABLA 4 DOCUMENTO DE ESPECIFICACIÓN DE REQUISITOS: SECCIONES 1 A 5.....	52
TABLA 5 DOCUMENTO DE ESPECIFICACIÓN DE REQUISITOS: SECCIÓN 6.....	54
TABLA 6 DOCUMENTO DE ESPECIFICACIÓN DE REQUISITOS: SECCIÓN 7.....	55
TABLA 7 EJEMPLO DE ONTOLOGÍAS POR CADA CATEGORÍA.....	60
TABLA 8 EJEMPLO DE COMPARACIÓN DE ONTOLOGÍAS DE INFORMACIÓN PERSONAL.....	61
TABLA 9 EJEMPLO DE SELECCIÓN DE ONTOLOGÍAS DE INFORMACIÓN PERSONAL...	62
TABLA 10 TÉCNICAS DE RECOLECCIÓN EXPLÍCITA DE DATOS	68
TABLA 11 TÉCNICAS DE RECOLECCIÓN IMPLÍCITA DE DATOS.....	69
TABLA 12 VENTAJAS DE UN FORMULARIO WEB.....	70
TABLA 13 GUÍAS PARA EL DISEÑO DE UN FORMULARIO WEB.....	72
TABLA 14 ATRIBUTOS DEL PERFIL DE USUARIO	73
TABLA 15 ATRIBUTOS DE LAS PREFERENCIAS DEL USUARIO.....	73
TABLA 16 VENTAJAS DE LAS REDES SOCIALES.....	76
TABLA 17 VENTAJAS DE LOS EVENTOS DE LA TELEVISIÓN DIGITAL.....	76
TABLA 18 ESPECIFICACIONES TÉCNICAS DE LAS REDES SOCIALES.....	78
TABLA 19 FORMATO DE RESPUESTA DE SOLICITUD DE LAS REDES SOCIALES.....	83
TABLA 20 VENTAJAS DE SERVICIOS REST	92
TABLA 21 MÉTODOS DE ACCESO REST AL MÓDULO MRS	92
TABLA 22 ESPECIFICACIÓN DE CASO DE USO: AUTENTIFICACIÓN DE USUARIO	101
TABLA 23 ESPECIFICACIÓN DE CASO DE USO: REGISTRO DE USUARIO MEDIANTE FORMULARIO WEB	102
TABLA 24 ESPECIFICACIÓN DE CASO DE USO: REGISTRO DE USUARIO MEDIANTE REDES SOCIALES	103
TABLA 25 ESPECIFICACIÓN DE CASO DE USO: REGISTRO DE USUARIO MEDIANTE FORMULARIO WEB Y REDES SOCIALES	104
TABLA 26 ESPECIFICACIÓN DE CASO DE: MODIFICACIÓN DE PREFERENCIAS DEL USUARIO	105
TABLA 27 VENTAJAS DEL REPOSITORIO VIRTUOSO	109
TABLA 28 TABLA COMPARATIVA, DIFERENTES SERVIDORES WEB/APLICACIONES ...	114
TABLA 29 VENTAJAS DEL SERVIDOR WEB APACHE	115



UNIVERSIDAD DE CUENCA

Cláusula de derechos de Autor

Yo, **JUAN JOSÉ SÁENZ PEÑAFIEL** autor de la tesis "MODELAMIENTO DEL PERFIL DE UN USUARIO USANDO TECNOLOGIAS SEMANTICAS", reconozco y acepto el derecho de la Universidad de Cuenca, en base al Art. 5 literal c) de su Reglamento de Propiedad Intelectual, de publicar este trabajo por cualquier medio conocido o por conocer, al ser este requisito para la obtención de mi título de Ingeniero de Sistemas. El uso que la Universidad de Cuenca hiciere de este trabajo, no implicará afección alguna de mis derechos morales o patrimoniales como autor/a

Cuenca, 23 de Septiembre del 2014

Juan José Sáenz P.

Juan José Sáenz Peñafiel

C.I: 0104555131



UNIVERSIDAD DE CUENCA

Cláusula de derechos de Autor

Yo, *SEBASTIÁN RODRIGO ROMÁN JARRÍN* autor de la tesis "MODELAMIENTO DEL PERFIL DE UN USUARIO USANDO TECNOLOGIAS SEMANTICAS", reconozco y acepto el derecho de la Universidad de Cuenca, en base al Art. 5 literal c) de su Reglamento de Propiedad Intelectual, de publicar este trabajo por cualquier medio conocido o por conocer, al ser este requisito para la obtención de mi título de Ingeniero de Sistemas. El uso que la Universidad de Cuenca hiciere de este trabajo, no implicará afección alguna de mis derechos morales o patrimoniales como autor/a

Cuenca, 23 de Septiembre del 2014

Sebastián Rodrigo Román Jarrín

C.I: 0104475785



UNIVERSIDAD DE CUENCA

Cláusula de Propiedad Intelectual

Yo, Sebastián Rodrigo Román Jarrín autor de la tesis "MODELAMIENTO DEL PERFIL DE UN USUARIO USANDO TECNOLOGIAS SEMANTICAS", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor.

Cuenca, 23 de Septiembre del 2014

Sebastián Rodrigo Román Jarrín

C.I: 0104475785



UNIVERSIDAD DE CUENCA

Cláusula de Propiedad Intelectual

Yo, Juan José Sáenz Peñafiel autor de la tesis "MODELAMIENTO DEL PERFIL DE UN USUARIO USANDO TECNOLOGIAS SEMANTICAS", certifico que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor.

Cuenca, 23 de Septiembre del 2014

Juan José Sáenz P.

Juan José Sáenz Peñafiel

C.I: 0104555131

AGRADECIMIENTOS

Agradecemos al equipo que conforma el proyecto de “Aplicación de Tecnologías Semánticas para Disminuir la Sobrecarga de Información en Usuarios de TV digital” por permitirnos ser parte del mismo y brindarnos su ayuda dentro del desarrollo de esta tesis. Principalmente al Ing. Mauricio Espinoza Mejía, MSc. PhD, quien además de ser director del proyecto fue nuestro tutor y mentor a lo largo del desarrollo de esta tesis y con sus conocimientos pudo ayudarnos a completarla.

DEDICATORIAS

Dedico este trabajo de tesis a mis padres Rodrigo y Jenny, que siempre se esforzaron en darme la mejor educación. A mis hermanas Dany y Gaby que fueron un soporte para seguir adelante. A Karina que siempre estuvo en mis momentos difíciles apoyándome y a mis profesores y amigos que también me ayudaron en el desarrollo de mi vida profesional.

SEBASTIÁN

Dedico este trabajo de tesis a mi madre Saida que siempre ha estado conmigo, buscando lo mejor para mí y ha sido el mejor apoyo dentro de mi carrera. A mi padre que me cuida día a día en mi caminar. A mi hermano Marcelo y todos mis familiares que siempre están conmigo. A Vane que es una parte muy importante de mi vida y un apoyo gigante y a mis amigos que están en todo momento.

JUAN JOSÉ



UNIVERSIDAD DE CUENCA

INTRODUCCIÓN

En el presente capítulo, se planteará la definición del proyecto, los alcances especificados y los objetivos que serán cumplidos al final del proyecto.

CAPÍTULO 1

INTRODUCCIÓN

El proyecto aprobado y financiado por la Dirección de Investigación de la Universidad de Cuenca (DIUC) denominado: “Aplicación de Tecnologías Semánticas para Disminuir la Sobrecarga de Información en Usuarios de TV digital”, pretende diseñar un sistema de recomendación para la programación televisiva que tome en consideración las preferencias del usuario. El proyecto puede ser dividido en dos grandes etapas como se observa en la Figura 1.

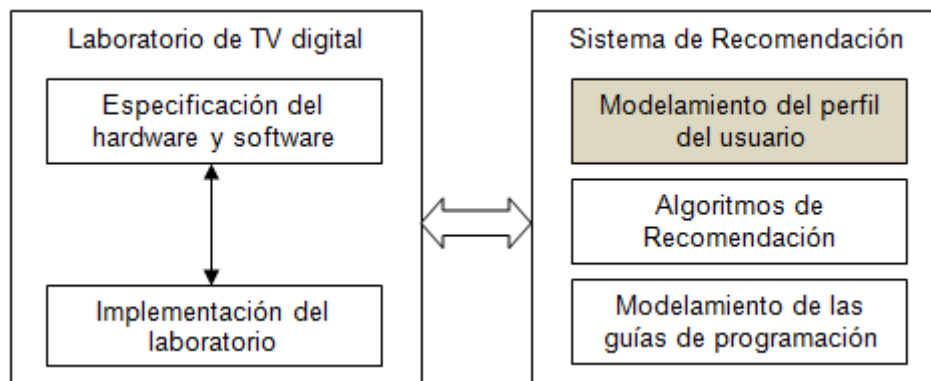


Figura 1 Esquema general del proyecto “Aplicación de Tecnologías Semánticas para Disminuir la Sobrecarga de Información en Usuarios de TV digital”

La etapa denominada “Laboratorio de TV digital” constituye la primera fase del proyecto. Esta etapa se encargará de buscar las alternativas óptimas para simular un escenario real de transmisión y recepción de señal televisiva. Este diseño estará formado por elementos de hardware y software que posibiliten transmitir y recibir múltiples contenidos televisivos. Por otra parte, la etapa denominada “Sistema de Recomendación”, se encargará de diseñar una infraestructura semántica por medio de la utilización de ontologías¹ para captar las preferencias de los televidentes y los contenidos de los programas. Esta etapa usará además algoritmos de recomendación que permitan seleccionar únicamente aquellos programas que son de interés para el televidente en base a sus preferencias, con el objetivo de no sobrecargarle de información.

¹ En esta tesis se usa el término ontología para referirse a la estructura de clasificación y las instancias dentro de una base de conocimientos.

El modelamiento del perfil del usuario usando tecnología semántica es el aspecto que cubrirá este proyecto de tesis. Este componente implementará las interfaces adecuadas para interactuar con los otros módulos del sistema de recomendación.

1.1. IDENTIFICACIÓN DEL PROBLEMA

El modelado de un usuario es una representación explícita de las propiedades de un usuario específico [1]. Es usado para razonar acerca de las necesidades, preferencias o comportamiento futuro de un usuario. Esta información resulta muy útil a la hora de diseñar nuevos sistemas, para saber lo que el usuario va a desear del mismo y poder orientarlo hacia sus necesidades y deseos.

La mayoría de los sistemas informáticos contienen algún tipo de modelado implícito de sus usuarios, pero dicho modelo en la mayoría de los casos, no puede variarse según el uso que el usuario dé al sistema y, por tanto, tampoco el sistema ofrece una respuesta adaptada.

De hecho, una interpretación incorrecta de un usuario lleva a decisiones equivocadas de adaptación, que pueden resultar en frustración, pérdida de confianza, disminución de la motivación para utilizar el sistema, etc.

El enfoque para modelar el perfil de un usuario que se utiliza en la mayoría de los sistemas de recomendación está basado en el comportamiento del usuario con el sistema [2]. Este tipo de representación usa comúnmente un modelo de clases binarias para definir lo que los usuarios encuentran interesante y poco interesante. Técnicas de aprendizaje automático son entonces usadas para encontrar elementos potenciales de interés respecto al modelo binario [3].

Uno de los problemas del modelamiento de perfil basado en un formato binario es que este tipo de representación no permite integrar con facilidad cualquier conocimiento de dominio que puede estar disponible. De hecho el formato para representar las distintas caracterizaciones de un usuario es un tema clave al momento de modelar el perfil de un usuario. Un simple documento legible por una persona y una lista de datos no son suficientes, dado que esta información no puede ser automatizado con facilidad y sobre todo no es accesible por una máquina.

En los últimos años, una serie de sistemas adaptables al usuario han estado explotando las ontologías a efectos de obtener una representación semántica,

una adquisición de conocimientos automática, y la creación de soluciones arquitectónicas interoperables y reutilizables (ver por ejemplo [4-6]). Las ontologías modelan conceptos y relaciones en un alto nivel de abstracción, proveyendo una semántica rica para el trabajo humano y el formalismo requerido para que los computadores ejecuten procesamiento mecánico y razonamiento.

En este trabajo exploramos un enfoque ontológico para el modelado de perfiles de usuario, el mismo que será usado en un sistema de recomendación. Específicamente, estamos interesados en un modelo formal que permita tanto a usuarios como a máquinas determinar las preferencias televisivas de un usuario.

1.2. JUSTIFICACIÓN

Con el fin de gestionar la enorme cantidad de información disponible en la actualidad la gente necesita apoyo específico de sistemas de recomendación inteligentes. Esta situación es una necesidad en diferentes campos de aplicación como: medios digitales, la Web, comercio electrónico, aprendizaje en línea, TV digital, etc. En todos estos campos de aplicación, la información destino depende de las necesidades y preferencias del usuario. Así, uno de los factores clave en los sistemas de recomendación es la captura y representación de las preferencias del usuario.

En la literatura podemos identificar una gran cantidad de propuestas para modelar ontológicamente el perfil de un usuario para diferentes aplicaciones como búsqueda en la Web [8, 9], gestión de información personal [10], o recomendación de noticias [11]. Inclusive, en el campo específico de esta tesis, existen varias propuestas que intentan modelar el perfil para un usuario de TV digital (ver ontología LUMO² por ejemplo). A pesar de estos esfuerzos, no podemos hablar todavía de una ontología estándar para estos propósitos, la mayoría de las ontologías creadas son específicas a la aplicación, cada una de ellas elaborada para un dominio particular.

La diversidad en cuanto a los datos que se consideran primordiales para modelar el perfil de un usuario nos obliga a estudiar diferentes ontologías de dominio que nos permita crear una ontología que permita describir de manera

² <http://mklab.iti.gr/project/lumo>
Román Jarrín Sebastián Rodrigo
Sáenz Peñafiel Juan José

adecuada el conocimiento relevante de un usuario de la TV digital. Para nuestros propósitos estamos considerando que el modelo abarque tanto conceptualizaciones de las características del usuario (tales como profesión, edad, género, experiencia, etc.), como sus intereses televisivos, los mismos que pueden cubrir entidades de diferentes dominios.

1.3. ALCANCE

De forma general el proceso de personalización y contextualización de las preferencias de un usuario puede ser dividido en tres aspectos:

1. La identificación de los datos a ser modelados y su formato de representación.
2. La determinación de las fuentes de información más apropiadas para la extracción de las preferencias.
3. El uso adecuado de este conocimiento para filtrar la información más relevante para el usuario.

El alcance de este trabajo está orientado hacia la búsqueda de una solución para los dos primeros aspectos. Es necesario hacer notar que la modelación del perfil del usuario planteada en este trabajo, tiene como meta final proveer una base de conocimiento para el sistema de recomendación de programación televisiva que se está implementado como parte del proyecto de investigación mencionado en la introducción. En dicho sistema se plantea que el uso de ontologías para modelar el perfil de un usuario permite que el comportamiento inicial del televidente sea emparejado con los conceptos existentes en la ontología. Este enfoque ontológico para modelar el perfil de un usuario ha demostrado ser eficaz en la solución del problema del arranque en frío en los sistemas de recomendación, ya que permite la propagación de un pequeño número de conceptos iniciales a otros conceptos relacionados explotando la estructura ontológica del dominio [12-14].

Por tanto, en este trabajo se explotará una representación semántica de las preferencias del usuario y se propondrá el uso de la Web social³ como fuente principal para extraer las preferencias de un usuario de TV digital. Nuestra hipótesis es que éste gran depósito, multifacético de datos sociales puede ser

³ El término Web social hace relación al cambio de paradigma que se ha producido en la Web pasando de un enfoque centrado en la máquina hacia un enfoque centrado en el usuario y en la comunidad

usado para extraer las preferencias implícitas de los usuarios de programas televisivos.

1.4. OBJETIVO GENERAL

Modelar formalmente las preferencias de un usuario, usando tecnología semánticas y recursos externos que permitan enriquecer esta información.

1.5. OBJETIVOS ESPECÍFICOS

- Identificar los aspectos que permitan modelar el perfil estático y dinámico de un usuario de TV digital.
- Analizar diferentes ontologías que permitan modelar el perfil de un usuario y extender éstos recursos en caso necesario.
- Analizar y seleccionar diferentes fuentes de información que permitan recolectar los datos estáticos y dinámicos de las preferencias de un usuario.
- Almacenar la información recolectada desde las fuentes seleccionadas en un formato accesible por humanos y máquinas.
- Desarrollar e Implementar la solución prototipo de modelamiento de usuario, la cual incluye: recolección del perfil, inferencia y consulta.

1.6. ESTRUCTURA DE LA TESIS

Esta tesis será dividida en seis capítulos tomados desde el capítulo actual que trata de definición del proyecto, alcance, objetivos y estructura de la tesis. A continuación se dará un breve detalle de los temas que serán tratados en cada uno de los capítulos siguientes:

En el capítulo 2 denominado Contexto Tecnológico se presentará el marco teórico y conceptual de los temas a ser tratados dentro del desarrollo de la tesis, presentando definiciones, elementos, tipos y justificaciones. Los temas a tratar son Perfiles de Usuario y su modelamiento, Redes Sociales, Ontologías y Metodologías para su Creación.

Continuando con el desarrollo de la tesis, en el capítulo 3 se detallará la metodología de creación de ontologías que se seleccionará dentro del contexto tecnológico en el capítulo 2, tomando en consideración los beneficios de haber seleccionado esa metodología. Los temas a tratar en el capítulo 3 serán el paso a paso de la construcción de una ontología con el método seleccionado.

Posteriormente, en el capítulo 4 se detallará el método usado para la recolección de datos de perfil de usuario desde fuentes externas, tomando en consideración los datos tomados de fuentes externas como son Redes Sociales y formulario de recolección de datos. El capítulo dará a detalle los procesos y mecanismos usados para la captura de los datos del perfil de usuario considerando los datos a ser obtenidos y las fuentes de donde serán recolectados.

Como siguiente capítulo, el capítulo 5 detallará todo el procedimiento seguido para la construcción del prototipo de recolección de datos de perfiles de usuario mediante tecnologías semánticas. Dentro de este capítulo se tratará los temas de la configuración de repositorio, la definición de arquitectura de la aplicación y el proceso de análisis, diseño, implementación y pruebas del prototipo definitivo, en el cual se tendrá los resultados obtenidos en el proyecto, tanto en análisis previo como en construcción de prototipo

Para finalizar con el capítulo 6 de conclusiones y recomendaciones. Adicional a eso en este capítulo se plantearán trabajos futuros dentro de la misma línea de orientación de perfiles de usuario y tecnologías semánticas.



UNIVERSIDAD DE CUENCA

CONTEXTO TECNOLÓGICO

En este capítulo, se presentará el contexto tecnológico de los temas principales a tratar dentro de los capítulos posteriores.

CAPÍTULO 2

CONTEXTO TECNOLÓGICO

En este capítulo se hará una revisión del contexto tecnológico que será utilizado dentro del modelamiento del perfil de usuario con tecnologías semánticas. Para lo cual primero se definirá lo que entendemos por perfil de usuario y sus formas de modelarlo. Luego se procederá a describir las redes sociales y su uso dentro del proyecto de tesis. Posteriormente, se describirá las diversas herramientas que serán utilizadas como parte del prototipo planteado, como lo son Ontologías, sus elementos y las metodologías de creación de ontologías.

2.1. PERFIL DE USUARIO

En esta sección aclararemos lo que es un perfil de usuario y sus formas de modelarlo, debido que al ser la base de esta tesis debe ser bien definido y entendido. Estos conceptos serán usados en el capítulo 3. Allí se utilizará el concepto de perfil de usuario que será modelado con tecnologías semánticas y su posterior instanciación con datos provenientes de diferentes fuentes de información.

2.1.1. Definición

“En sentido general, un perfil de usuario es un conjunto de datos que se refieren al usuario de un servicio informático. Un modelo de usuario es una fuente de conocimientos que contiene adquisiciones sobre todos los aspectos del usuario que pueden ser útiles para el comportamiento del sistema.”[15]

“En el contexto de la informática, un usuario es una persona aquel que utiliza un sistema informático. Para que los usuarios puedan obtener seguridad, acceso al sistema, administración de recursos, etc., dichos usuarios deberán identificarse. Para que uno pueda identificarse, el usuario necesita una cuenta (una cuenta de usuario) y un usuario, en la mayoría de los casos asociados a una contraseña. Los usuarios utilizan una interfaz de usuario para acceder a

los sistemas, el proceso de identificación es conocido como identificación de usuario o acceso del usuario al sistema (del inglés: "log in").". [16]

Estas dos definiciones permiten delimitar las acciones de un perfil de usuario, el cual puede ser usado para personalizar un sistema en específico, basado en las preferencias y características del usuario. Esto permite que el sistema se adapte a un usuario en específico. Algunos ejemplos de datos que pueden ser parte de un perfil de usuario pueden ser: edad, sexo, gustos específicos de un usuario, entre otros; que hacen que un usuario maneje su propio perfil.

Este tipo de datos especificados como ejemplo buscan pasar de un sistema generalizado a un sistema específico basado en parámetros de preferencias de un usuario, lo cual genera que una aplicación sea más funcional, debido a que personaliza a la aplicación de acuerdo a lo que el usuario incluye como información propia. Por ejemplo, si un usuario usa un sistema de recomendación televisiva que le pide su edad, el rango de recomendación del sistema buscará programas acordes a la edad ingresada, que sean permitidos para su edad y que además sean llamativos de acuerdo a estadísticas sobre lo que personas de ese rango de edad miran en la televisión.

2.1.2. Información de usuario

El modelamiento de perfiles de usuario se puede realizar principalmente con técnicas de extracción de información implícita y explícita [17]. En esta sección se describirá a continuación las técnicas usadas comúnmente en la literatura para la captura de datos para un perfil de usuario. Una vez descritas las diferentes técnicas, se procederá a seleccionar el método más adecuado para los fines que persigue este trabajo de tesis

2.1.2.1. Extracción implícita

*"Datos implícitos son datos que no se proporcionaron intencionalmente, pero obtenidos de los flujos de datos disponibles, ya sea directamente o por medio de análisis de datos explícitos. Por ejemplo a través de encuestas y formularios de registro de miembros."*⁴

Una primera forma de construir el perfil de un usuario, es usar una técnica de modelamiento implícito, que como nos describe la definición toma datos desde la misma aplicación sin que el usuario tenga que ingresarlos. El principal objetivo de este modelamiento es alimentar siempre el perfil (desde la

⁴ <http://whatis.techtarget.com/definition/implicit-data>

aplicación) con datos reales sin necesidad que el usuario actualice estos datos a cada momento. En el caso concreto de esta tesis, se busca que la interacción del usuario con el televisor permita construir de manera implícita los datos del perfil, junto con los datos respectivos de las redes sociales.

- La principal ventaja de este modelamiento es que el usuario puede tener su perfil actualizado sin mucho esfuerzo.
- La principal desventaja es la dificultad de recolectar adecuadamente la información que será parte del perfil.

2.1.2.2. Extracción explícita

“Datos explícitos son datos que se proporcionan intencionalmente tomando su significado directo sin ser analizado ni interpretado. Ejemplos de datos explícitos incluyen información proporcionada en las encuestas y las solicitudes de afiliación.”⁵

El segundo método de modelamiento recibe el nombre de explícito y se basa, como dice la definición, en solicitar al usuario ingresar los datos que necesita el perfil, de manera que el usuario sea quien actualice la información necesaria. Un inconveniente de este método es que los datos ingresados son proclives a errores o aciertos al momento de ingresar la información. Los datos recolectados son solicitados al usuario comúnmente mediante el uso de formularios de datos personales y de selección de preferencias de la persona que va a llenar su perfil.

- La ventaja principal que se tiene es la facilidad para el desarrollador de la toma de datos.
- Como principal desventaja se considera los posibles errores de ingreso de datos por parte del usuario y además el tiempo que le toma ingresar los mismos. El ingreso de datos por parte del usuario suele generar poco interés en la aplicación por parte del usuario.

2.1.2.3. Extracción híbrida

Como resultado de la unión de los métodos antes mencionados se genera un método que adquiere parte de cada uno y busca optimizarlas, se lo denomina *híbrido*. Este método captura información de los usuarios usando datos explícitos como datos base y posteriormente los alimenta con datos implícitos capturados por la aplicación. El método híbrido maneja la ventaja principal de optimización de datos gracias a los datos obtenidos directamente del usuario y además maneja datos constantemente actualizados por la aplicación

⁵ <http://what-is.techtarget.com/definition/explicit-data>

Para el proyecto presentado en esta tesis el método a usar será el método híbrido. En una primera instancia se recogerá explícitamente datos base como edad y sexo del usuario. En una segunda instancia se enriquecerá los datos del perfil usando métodos implícitos tomados de la interacción del usuario con la televisión y redes sociales.

En el capítulo 4 se describirá el proceso tecnológico de recolección de datos desde fuentes externas, se utilizarán los conceptos de los métodos de modelamiento de datos.

2.1.3. Modelamiento de perfiles de usuario

Una vez definidos los diferentes métodos para la captura de los datos del perfil de un usuario, se puede proceder al modelamiento del perfil con dichos datos. Para modelar perfiles de usuario es necesario diseñar un modelo en el que se incluya los gustos de un usuario en específico tomando en consideración el dominio y el entorno donde se usará el perfil. En la literatura se pueden identificar al menos cuatro métodos para realizar el modelado de un perfil. Estos métodos para modelar un perfil, serán presentados a continuación.

2.1.3.1. Modelado del perfil con preferencias explícitas

Este método de modelamiento busca generar un perfil de usuario basado en la información explícita que pueda obtenerse a partir de los datos que tenga un usuario dentro de un sitio web o un servicio web. La información que usualmente puede ser modelada mediante este método son: datos demográficos como edad, sexo, etc. y los datos profesionales del usuario, como títulos, trabajo, etc. Además, es posible modelar datos sobre gustos, aficiones, política, etc. Cabe resaltar que esta información en algún momento debe ser revisada por el usuario debido a que es obtenida de manera automática.

Dentro del modelamiento con preferencias explícitas surge el mayor inconveniente de que un usuario no siempre está dispuesto a tener toda su información dentro de un sitio por diferentes razones, por ejemplo seguridad; por lo que el método de modelamiento explícito genera un problema que debe ser evaluado en su desarrollo.

2.1.3.2. Modelado del perfil basado en contenidos

Para el método de modelamiento basado en contenidos se aprovecha los eventos y publicaciones generadas por un usuario para analizarlos y modelarlos como parte del perfil. Es decir, el modelamiento se basará en

características de los contenidos generados por el usuario, para posteriormente realizar un mapeo de estas características con aquellas características necesarias en la aplicación.

Generalmente los sistemas que usan este modelado manejan dos componentes principales, el primero un analizador de contenido que es el encargado de la extracción y análisis del contenido del usuario y el segundo, un mecanismo de aprendizaje que requiere técnicas de aprendizaje automático para generalizar los datos recogidos del contenido publicado del usuario.

Además, dentro de los componentes antes mencionados y para el modelamiento de un perfil se utilizan técnicas de minería de texto (análisis palabra a palabra), retroalimentación de contenido entre la aplicación y el usuario, análisis estadístico de uso de términos para saber los términos más comunes y categorizarlos, entre otras técnicas.

2.1.3.3. Modelamiento del perfil basado en la interacción entre pares

Para el modelamiento entre pares se toma en consideración la interacción del usuario dentro de un entorno como medios sociales o dentro del sistema mismo, es decir se tomará el manejo de interacciones que sean útiles al modelo del perfil necesario pero tomándola entre pares. Estas interacciones entre pares buscan dar como resultado una categorización de los datos de usuarios basados en resultados similares de los pares analizados.

Para el modelamiento del perfil basado en interacción entre pares existen dos técnicas principales:

Basado en colaboración:

Esta técnica es la más utilizada dentro del modelamiento entre pares, la idea principal de este modelamiento es un análisis entre pares del comportamiento de grupos de usuarios, para luego extraer información significativa que este dentro de los grandes volúmenes de información obtenida sobre el comportamiento y las preferencias de la mayor cantidad de usuarios posibles.

Minería de red social:

Esta técnica tiene como base las redes sociales y se basa en entender el comportamiento de conexiones de un usuario dentro de la red social, es decir analizar las conexiones sociales entre usuarios y sus intereses comunes. Para ello se realiza una minería de datos basado en gráficas de comportamiento y de interacción entre usuarios. Además, mediante la

minería de datos en la red social se pueden obtener datos demográficos de un usuario, sus preferencias y expectativas dentro de la red social analizada, con lo que se puede completar un modelo de perfil.

2.1.3.4. Modelamiento del perfil basado en el conocimiento

Este modelamiento tiene como objetivo solucionar el problema del denominado arranque en frío, es decir poder brindar recomendaciones a pesar de la falta información. Para solucionar este problema se busca la utilización del conocimiento previo generado o inferido a través de tecnologías semánticas. Una forma común de modelar la información del perfil basado en conocimiento, es hacer uso de ontologías estructuradas previamente y que pueden ser aplicadas al dominio. Para el uso de un modelamiento basado en conocimiento se requiere de expertos en el área que generen modelos basados en el conocimiento previo.

2.1.3.5. Modelamiento Híbrido de perfil

El modelamiento híbrido combina el modelamiento basado en el conocimiento y el modelamiento basado en el contenido. Combinando estos dos métodos se busca sanar la falta de información obtenida en un método basado en el contenido y explotar el conocimiento de la mejor manera tomando el contenido como base. La idea principal de un método híbrido es complementar la información obtenida en un perfil de usuario mediante conocimiento de un modelado ontológico que complemente las deficiencias en la falta de contenidos a través de inferencia que pueda completar esa información faltante.

Considerando estas características del modelamiento híbrido y los requerimientos de esta tesis, este tipo de modelamiento será el utilizado dentro del desarrollo del prototipo estipulado dentro de los objetivos planteados en el capítulo 1.

2.1.4. Creación de un modelo perfil de usuario

Existen diversos criterios de diseño a la hora de modelar un perfil de usuario, el criterio de diseño Gruber [18] propone que los datos/atributos del perfil de usuario deben brindar claridad, coherencia, extensibilidad, un mínimo de sesgo para codificación, y un mínimo compromiso ontológico. Estos aspectos se deben tomar en cuenta a la hora del proceso de creación del perfil. A continuación se especifican algunas pautas para la creación del modelo de perfil de usuario. [19]



La información en los modelos de perfiles no deben contener conceptos de alto nivel, sino más bien instancias de conceptos y nombres de atributos lo más simple posible En la figura 2 se muestra cómo se modela cierta información del perfil de un usuario ICQ⁶. La idea es identificar primero las categorías de datos que requieren ser modelados y luego establecer las relaciones necesarias entre éstas categorías y los términos del modelo ontológico usado para representar el perfil

ICQ Profile Category	ICQ Profile Property	Modeling in the Profile Ontology
Home	Street Address	Slots of the “Living Conditions” class
	Zip	
	City	
	State	
	Country	
Place of Birth	City	Slots of the “Person” class
	State	
	Country	
Personal Info	Homepage	Slot of the “Person” class
	Gender	Slot of the “Person” class
	Age	Not necessary, may be calculated by the date of birth
	Date of Birth	Slot of the “Person” class
	Zodiac Sign	Not necessary, may be calculated by the date of birth
	Spoken Languages	Could be added as Instances of the “Education” class

Figura 2 Modelo de un perfil de usuario en ICQ [50]

Los modelos de perfil de usuario consideran conceptos apropiados para adaptar e incluir en la ontología. Tazari [20] sugiere los siguientes conceptos importantes para el perfil de usuario: identidad del usuario, características, capacidades, preferencias generales, estado del usuario, y preferencias específicas. Otros conceptos como actividades actuales, ubicación, estado de movimiento y orientación se mencionan además como importantes para definir el perfil. También el autor propone un grupo de parámetros relativos a la información personal (nombre, fecha de nacimiento, dirección, cuenta bancaria y tarjeta de crédito), características generales (factores físicos: peso y talla, discapacidades físicas y habilidades: leer, hablar y escribir en otro idioma),

⁶ ICQ, <http://www.icq.com/>
Román Jarrín Sebastián Rodrigo
Sáenz Peñafiel Juan José

educación, ocupación, experiencia y estado de usuario; todos estos conceptos engloban el perfil básico para un usuario en general.

Los Intereses ([21], [22], [23], [24]) y preferencias ([25], [24]) se consideran de especial importancia para la mayoría de aplicaciones que incorporan perfiles de usuario. Los intereses son, en algunos casos organizados en jerarquías de conceptos [25], [21]. Las habilidades del usuario, tanto física y mental también parecen ser relevantes [26]. El factor de género también se ha demostrado que influyen en el comportamiento de los diferentes usuarios mientras interactúan con un mismo sistema [25].

La experiencia del usuario, ya sea relacionado con la informática o relacionado con otro dominio es un concepto necesario para muchos perfiles de aplicaciones [24].

La mayor parte de las recomendaciones identificadas en la literatura sobre que atributos debería contener un perfil de usuario, han sido consideradas en el modelamiento del perfil implementado en esta tesis. La descripción del modelo del perfil de usuario usado en este proyecto se detallará en el *capítulo 3*.

2.2. REDES SOCIALES

Dentro del contexto a desarrollar en este proyecto de tesis, las redes sociales juegan un papel fundamental dentro del proyecto. Las redes sociales son en nuestro caso la fuente de alimentación de datos principal por parte del usuario para complementar el perfil con datos reales sin necesidad de que el usuario deba ingresarlos.

2.2.1. Definición

“Una red social es una forma de representar una estructura social, asignándole un grafo, si dos elementos del conjunto de actores (tales como individuos u organizaciones) están relacionados de acuerdo a algún criterio (relación profesional, amistad, parentesco, etc.) entonces se construye una línea que conecta los nodos que representan a dichos elementos.”⁷

Tomando el concepto de red social y relacionándolo a la informática, tenemos

⁷ http://es.wikipedia.org/wiki/Red_social

el concepto de *servicio de red social* que se definirá a continuación:

“Un servicio de red social es un medio de comunicación social que se centra en encontrar gente para relacionarse en línea. Están formadas por personas que comparten alguna relación, principalmente de amistad, mantienen intereses y actividades en común, o están interesados en explorar los intereses y las actividades de otros.”⁸

Para hablar de redes sociales en internet o servicio de Red social, debemos hablar de aplicaciones en la web orientada a la conexión social y basadas en el concepto de red de nodos de usuarios que interactúan de diversas maneras entre sí, simulando una conexión social entre todos los usuarios conectados.

Dichas redes de conexión generan interacción y comunicación entre usuarios y gracias a estos parámetros generan cantidades de información muy grandes, que pueden ser usadas en muchas aplicaciones que necesitan ser enriquecidas con datos valiosos.

Los datos dentro de una red social dependen exclusivamente de la información que el usuario ingrese dentro de la misma. En algunos casos los volúmenes de datos que se puede manejar son demasiado extensos para manejarlos en su totalidad. Debido a esto, es necesario filtrar y transformar la información útil para las aplicaciones que consuman posteriormente estos datos generados por él.

2.2.2. Selección de redes sociales

En la actualidad las redes sociales toman un papel muy importante en el intercambio de información, es por eso que es preciso valerse de estas tecnologías para alimentar nuestra base de conocimiento, pues como es evidente las redes sociales pueden proveer información implícita del usuario debido a la cantidad de información que tiene dentro de sí. Por ejemplo un usuario, registrado en la red social Twitter, puede contener los datos de las cuentas que sigue el usuario y con esas cuentas se puede deducir algunos gustos del usuario. De igual manera en la red social Facebook, tomando en consideración las páginas que le gustan a un usuario específico se puede determinar una inferencia de gustos por tipo de páginas que ha seleccionado como “Me gusta”.

⁸ http://es.wikipedia.org/wiki/Servicio_de_red_social

Debido a la existencia de decenas de redes sociales⁹, es necesario establecer algunos parámetros que permitan identificar aquellas que son más relevantes dentro del proyecto planteado en la tesis. Algunos criterios han sido considerados para identificar las redes sociales más relevantes: cantidad de usuarios registrados, medios de acceso y las diferentes alternativas para acceder a ellas, además de su disponibilidad dentro de nuestra región y particularmente en **Ecuador**.

La lista que se detalla en la tabla **1** muestra la información de cada red social indagada, comparación de ventajas y desventajas, además su estado (SI, para el caso que sea factible utilizar, caso contrario NO para las redes que no vayan a ser usadas) y comentarios que determinan fundamentalmente un motivo de decisión en la toma de la red social para el proyecto de tesis.

RED SOCIAL	VENTAJAS	DESVENTAJAS	ESTADO	COMENTARIOS
Facebook	Segunda Red Social con más registro de usuarios. Contiene API ¹⁰ de desarrollo.	Privacidad de contenido	Si	Es la segunda red social con más usuarios registrados en el mundo. ¹¹ Además Facebook cuenta con más de 7 millones de usuarios registrados en Ecuador ¹² por lo tanto es imprescindible para obtener información.
Twitter	Red Social entre las 15 con más registro de usuarios. Contiene API de desarrollo.	Dificultad de tomar la información de cada tweet debido a que son escritos en lenguaje natural y la manera de escribir varia de usuario a usuario.	Si	Twitter va ganando espacio en el Ecuador, por lo que genera grandes cantidades de información, lastimosamente cada publicación de información es diferente por lo que genera dificultad de procesamiento.
Google+	Contiene API, Red social para subir y	Ha ganado espacio	Si	Google+ se compone de YouTube, contiene información

⁹ http://en.wikipedia.org/wiki/List_of_social_networking_websites

¹⁰ Application Program Interface. Plataforma de desarrollo de aplicaciones dentro de la red social.

¹¹ http://en.wikipedia.org/wiki/List_of_social_networking_websites

¹² <http://www.owloo.com/facebook-stats/ecuador/>

	compartir vídeos La Red Social con más registro de usuarios.	lentamente en Ecuador por su similitud de contenidos con Facebook		de videos compartidos.
Deezer	Contiene API, sitio web que ofrece música de forma gratuita e ilimitada	Su conexión generalmente las personas la hacen mediante inicio de sesión Facebook	No	Esta red social no es muy utilizada en Ecuador, pero contiene una api muy intuitiva y podría llegar a ser útil la información
Orkut	Red social más popular en India y Brasil. Contiene API de desarrollo.	Escasos usuarios en Ecuador	No	Red social principalmente de Brasil, no tiene mayor influencia en Ecuador
Netflix	Capacidad para revisar las películas/series que ve el usuario	Red de pago lo que limita la cantidad de usuarios en Ecuador.	No	No se puede extraer información de usuario
LinkedIn	Datos profesionales de las personas	Datos poco influyentes para películas y series	No	LinkedIn proporciona datos útiles para el campo profesional y no de entretenimiento como la televisión digital.
Flixster	Red social especializada en películas	La red social no está disponible en la zona.	No	No disponible en el Ecuador

Tabla 1 Tabla comparativa de las redes sociales más importantes en el medio

Como se detalló en la tabla **1** las redes sociales a ser utilizadas en el proyecto de tesis serán Facebook, Twitter y Google+ principalmente por que cumplen con los parámetros antes mencionados. Las tres redes sociales alcanzan una cantidad importante de usuarios registrados¹³, además tienen una plataforma de desarrollo y están disponibles en nuestro país.

En el capítulo 3 se detallará el uso de las redes sociales como fuente de datos que serán usados en la población de un perfil de usuario de televisión digital.

¹³ http://en.wikipedia.org/wiki/List_of_social_networking_websites

2.3. ONTOLOGÍAS

Dentro del modelamiento del perfil de usuarios se busca la manera de almacenar los datos obtenidos del usuario y sus interacciones mediante un mecanismo diferente al tradicional mediante bases de datos, generando la apertura a la tecnología semántica por medio de las ontologías.

Para definir una ontología primero proponemos varios conceptos dados por diversos autores, con el objetivo de aclarar este concepto:

- *"Una ontología es un vocabulario acerca de un dominio: términos + relaciones + reglas de combinación para extender el vocabulario".* Neches[27]
- *"Una ontología es la especificación de una conceptualización".* Gruber [28].
- *"Una ontología es un conjunto de términos estructurados que describen un tema o 'dominio'. Su propósito es proporcionar una estructura básica para una base de conocimientos".* Swartout[29]

La ontología define rigurosos esquemas conceptuales dentro de uno o varios dominios dados; con la finalidad de facilitar la comunicación y el intercambio de información entre diferentes sistemas y entidades. Su uso se extiende desde la web semántica al dominio semántico en la inteligencia artificial y la gestión del conocimiento, donde la ontología se relaciona con vocabularios fijos. También existen ciertos casos donde para algunos dominios de problemas, se deben crear esquemas más especializados para convertir en útiles los datos.

Para ejemplificar ontologías que buscan modelar perfiles de usuario, podemos nombrar dos ontologías que sirven de referencia como lo son FOAF (Friend of a Friend)¹⁴ y GUMO (General User Model Ontology) [30]. Ambas ontologías de alto nivel buscan modelar un perfil de usuario con una representación de datos básicos como nombre, edad, ocupación, etc. Además las ontologías antes mencionadas proveen características concretas de representación de un perfil que incluye intereses de usuario, por lo que se hacen muy completas a la hora de modelar un perfil.

Dentro del contexto de televisión digital que trata esta tesis, cabe mencionar el proyecto LinkedTV¹⁵ que es realizado en Europa y utiliza tecnologías

¹⁴ <http://xmlns.com/foaf/spec/>

¹⁵ <http://www.linkedtv.eu/>

semánticas para el desarrollo de soluciones de integración entre televisión y contenido web disponible, de manera que genera contenido enriquecido al usuario. Como resultado de este proyecto se ha generado la ontología LUMO (Linked User Model Ontology) ¹⁶ que modela un perfil de usuario semánticamente. Esta y otras ontologías han sido usadas para modelar el perfil de un usuario de TV digital. La información sobre la creación del modelo ontológico será descrita en el capítulo 3.

2.3.1. Ontologías y su representación del conocimiento

En la actualidad existen grandes cantidades de información que están disponibles en la Web y en diferentes fuentes, el reto importante es encontrar la manera de que esta información pueda utilizarse para la personalización y contextualización de un modelo de usuario. Para comenzar debemos deducir que una *ontología* se compone de su vocabulario y su base de conocimiento como estructura principal del modelo de usuario. Este modelo puede contener todos los dominios y/o conceptos específicos del usuario, también contener las relaciones más importantes a modelar. Dicha ontología también puede proporcionar uniformidad, conceptos compactos para el conocimiento confuso, sinónimos y una base de conocimiento multilinguaje para una inferencia más eficaz. [31]

En cuanto a los lenguajes de representación del conocimiento, que puedan expresar adecuadamente una ontología que incluya información relativa de las preferencias del usuario, existen un gran número de normas que se han propuesto por el Consorcio WWW (World Wide Web). Favorablemente aparece el modelo de datos **RDF**¹⁷ que sirve de base para este tipo de lenguajes ya que cubre las relaciones básicas de la semántica utilizando tripletas. Dentro del esquema de RDF/S las ontologías primitivas básicas se pueden expresar, lo que permite la inferencia de nuevas tripletas.

Las normas de serialización utilizado para RDF son por ejemplo XML¹⁸, N3 o Turtle. Muchos lenguajes de representación soportan esquemas semánticos extendidos basados en RDF, por ejemplo, SKOS¹⁹, el cual expresa la información de forma jerárquica. En lo que respecta a la representación del

¹⁶ <http://mklab.iti.gr/project/lumo>

¹⁷ <http://www.w3.org/TR/rdf-schema/>

¹⁸ <http://www.w3.org/XML/>

¹⁹ <http://www.w3.org/2001/sw/wiki/SKOS>

conocimiento formal, lenguajes como OWL²⁰ y su fundamento lógico se han convertido en opciones que prevalecen para la representación del conocimiento.

Existen varios lenguajes de ontología que se pueden utilizar como base para el vocabulario del perfil de usuario para la televisión digital; y también muchas ontologías que pueden servir como la base de conocimientos que describe el modelo de usuario permitiente para un sistema de personalización semántica. En esencia, la ontología debería proporcionar la estructura y las relaciones necesarias para formular y gestionar el modelo de usuario, y debe tener el conocimiento necesario para recuperar información adicional relevante a este modelo.

Las ontologías y lenguajes de ontología mencionadas previamente sirven como base para plantear el modelo ontológico de perfil de usuario descrito en el capítulo 3 del presente documento.

2.3.2. Elementos de una ontología

Para definir los elementos de una Ontología vamos a ilustrar con un ejemplo y a continuación se irá describiendo cada elemento de la ontología. (Ver figura 3).

“Instrumentos Musicales” se define como un CONCEPTO ya que es la idea principal a ser modelada y representa a una clase a ser modelada. Un concepto puede ser una clase, un objeto, una estrategia o un método.

²⁰ <http://www.w3.org/2001/sw/wiki/OWL>

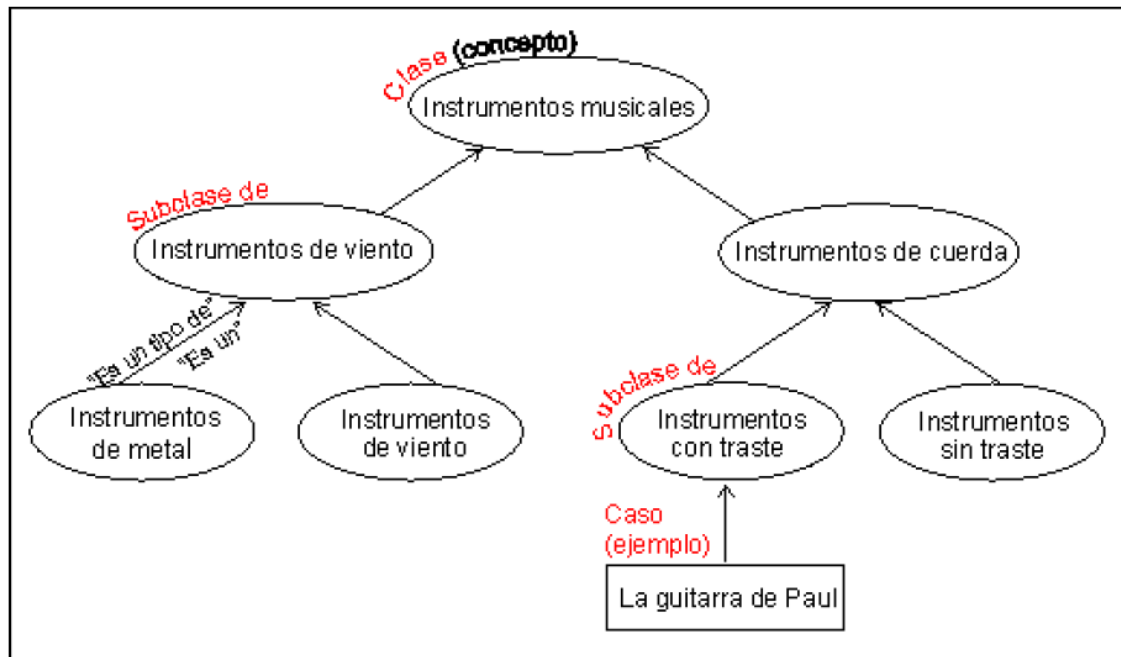


Figura 3 Ejemplo de Ontología²¹

“Marca” representa una PROPIEDAD del concepto, es decir un vehículo a motor tiene una marca, marca es un elemento que puede tener el CONCEPTO.

“Instrumentos de Viento” e “Instrumentos de cuerda” son una SUBCLASE de instrumentos musicales, pero también son CONCEPTOS pero que pueden ser especializados como subclases.

“Instrumentos de Metal” es un CONCEPTO relacionado a Instrumentos de Viento pero tipo de la misma, es decir un Instrumento de Metal es un tipo de Instrumento de viento o es un instrumento de viento. La RELACIÓN “es un tipo de” o también “es parte de” genera una relación de subclase y permite definir entidades complejas a partir de otras entidades más sencillas de manera incremental.

“La guitarra de Paul” es una INSTANCIA del concepto, es decir ya es un objeto específico, ya no es una generalización, es un objeto determinado.

2.3.3. Metodologías de creación de ontologías

Al igual que ocurre en la ingeniería de software cuando se desea crear un

²¹ <http://www.infor.uva.es/~sblanco/Tesis/Ontolog%C3%ADas.pdf>

nuevo producto, es necesario seguir una metodología para la construcción del producto; esta misma situación debe contemplarse para el proceso de creación de ontologías. El campo de estudios encargado de cubrir estas actividades se denomina ingeniería ontológica que es definida de la siguiente manera:

“La ingeniería Ontológica se refiere al conjunto de actividades que se refieren al proceso de desarrollo de una ontología, el ciclo de vida de una ontología, así como a las metodologías, herramientas y lenguajes necesarios para la construcción de ontologías.”²²

El concepto de ingeniería ontológica nos describe un proceso de construcción y mantenimiento del ciclo de vida de una ontología. Al igual que la ingeniería de software, la ingeniería ontológica requiere que este proceso sea ordenado y coherente para finalizar con un producto de calidad.

Existen diversas metodologías para la construcción de ontologías, pero por la cantidad de metodologías existentes a continuación se procederá a describir algunas de ellas.

2.3.3.1. CYC

La metodología CYC es una de las más antiguas debido a que fue publicada en 1990[32]. Esta metodología es muy general y simplifica su proceso en dos pasos:

- **Extracción manual del conocimiento común.-** En este paso la metodología busca extraer el conocimiento a ser modelado dentro de la ontología.
- **Utilización de una herramienta de procesamiento de lenguaje natural para la adquisición de nuevo conocimiento.-** El objetivo de este paso de la metodología es completar al conocimiento capturado mediante el procesamiento del lenguaje natural mediante una herramienta computacional que agregue conocimiento a la ontología.

2.3.3.2. USCHOLD Y KING

La metodología de Uschold y King [33], publicada posteriormente a la metodología CYC se basa en la construcción de ontologías basada en la reutilización de otras existentes. Para completar el proceso de creación, la metodología propone cinco pasos que se detallarán a continuación:

- **Identificación del propósito para el cual se construye la ontología.-** En este paso, la metodología busca definir el objetivo principal de la

²² <http://mayor2.dia.fi.upm.es/oeg-upm/index.php/es/researchareas/2-ontologicalengineering>

ontología, es decir identificar la necesidad principal de la construcción de la ontología.

- **Capturar los conceptos y las relaciones entre ellos.-** Identificar todos los conceptos y relaciones que tenga la ontología, es decir, haciendo una analogía al desarrollo de software, se busca las clases y las relaciones entre clases.
- **Codificación de la ontología.-** El paso de codificación es la construcción de la ontología como tal, es decir la implementación.
- **Evaluación de la ontología.-** Es el proceso de pruebas y evaluación de la ontología, por lo tanto es la verificación del funcionamiento de la ontología para el objetivo planteado.
- **Documentación de la ontología.-** El paso de documentación es el paso final cuando la ontología ha sido probada, en este paso se busca la documentación correcta para los futuros usos de la ontología.

2.3.3.3. GRÜNINGER Y FOX

La metodología de Grüninger y Fox [34] se basa en la construcción de ontologías considerando el uso que se dará a estas herramientas, es decir se analiza las preguntas que debe responder la ontología luego de su construcción. Los pasos que sigue esta metodología son los siguientes:

- **Definición de los escenarios motivadores.-** En este paso se define las posibles aplicaciones de la ontología, lo cual definirá los requisitos que debe cumplir la ontología.
- **Formulación de preguntas en lenguaje natural.-** En este paso se harán las preguntas que debe responder la ontología para poder definir el ámbito de uso de la ontología.
- **Especificación de la terminología.-** De las preguntas obtenidas en el paso anterior se busca obtener los conceptos principales, relaciones, propiedades, conceptos secundarios, etc., que sean necesarios para cumplir con las preguntas planteadas.
- **Formalización de las interrogantes.-** Una vez determinadas los elementos de la ontología se debe formalizar conceptos y desarrollarlos, es decir jerarquizar los conceptos de acuerdo a la necesidad.
- **Especificación de axiomas formales.-** Con los conceptos formalizados y desarrollados en el paso previo, la especificación de axiomas formales busca encontrar conceptos que llegan a ser claramente visibles sin explicación, por ejemplo el concepto persona que con su nombre son descritos sus propiedades directamente.
- **Verificación de la ontología.-** Proceso de pruebas para la ontología.

2.3.3.4. KACTUS

La metodología KACTUS [35] se basa en la abstracción de conceptos tomados de una base de conocimiento previamente definida. Para ello utiliza el siguiente proceso:

- **Especificación del contexto de la aplicación y el punto de vista de modelado.-** Se debe definir el dominio que tendrá la ontología, los objetivos que va a cumplir y los objetos que va a modelar. Además en este paso se debe definir si se hará una ontología estática-dinámica (con posibilidad de expansión) o un modelo funcional-casual (para cubrir una sola necesidad).
- **Realizar un diseño preparatorio en base a una ontología existente.-** En esta parte de la metodología se busca ontologías similares que puedan servir para ejecutar un proceso de reingeniería.
- **Refinamiento y estructuración de la ontología.-** Una vez obtenida una ontología similar se realiza la depuración de la misma tomando en consideración las necesidades de la ontología que se está desarrollando.
- **Documentación y reutilización de la ontología.-** En esta etapa se ejecuta el proceso de documentación final de la ontología para poder describir su uso y sus usos posteriores.

2.3.3.5. METHONTOLOGY

La metodología Methontology [36] tiene como característica que puede ser usada para crear ontologías iniciando desde cero y también para crear ontologías reusando las existentes. Esta metodología basa su proceso en las metodologías descritas en el campo de la Ingeniería de Software y consta de los siguientes pasos:

- **Especificación.-** Es la definición del alcance y además de la granularidad que tendrá la ontología.
- **Conceptualización.-** En este paso se realiza la estructuración y organización de los datos para la construcción del modelo ontológico, para ello se usa diagramas, tablas, etc.
- **Implementación.-** Este paso describe la construcción de la ontología en un lenguaje como RDF, OWL, entre otros, basado en la conceptualización antes obtenida.
- **Evaluación.-** En este paso se ejecuta un proceso de evaluación y pruebas de la ontología para soportar un correcto funcionamiento.

2.3.3.6. NeOn

La metodología NeOn [37] se basa en la creación de ontologías tomando en

consideración 9 escenarios diferentes. Estos escenarios intentan soportar la diversidad de posibilidades que se puede encontrar para modelar una nueva ontología.

- **Desde la especificación a la implementación.-** En este escenario se busca una implementación desde cero, tomando requisitos y modelándolos para posteriormente construir la ontología basado en esos requisitos.
- **Reutilizando y realizando reingeniería de recursos no ontológicos.-** Se busca la reingeniería como proceso de desarrollo, es decir primero se buscará recursos no ontológicos que se adapten a la necesidad y se los implementará dentro de un proceso de construcción de ontologías con recursos externos.
- **Reutilizando recursos ontológicos.-** Dentro del escenario propuesto se buscará reutilizar recursos ontológicos existentes y su posterior adaptación sin cambios para la necesidad pactada.
- **Reutilizando y realizando reingeniería de recursos ontológicos.-** En este escenario se buscará recursos ontológicos existentes y posteriormente se realizará reingeniería para reutilizarlos dentro del proyecto, de manera que se los adaptará cambiando los parámetros necesarios.
- **Reutilizando y fusionando recursos ontológicos.-** El proceso en este escenario se da cuando en la búsqueda de recursos ontológicos se encuentran varios que cubren la necesidad del proyecto, entonces se fusionará estos recursos para obtener un producto final.
- **Reutilizando, fusionando y ejecutando reingeniería de recursos ontológicos.-** Este escenario es muy similar al anterior con la diferencia de que el desarrollador de la ontología busca reorganizar los recursos mediante reingeniería bajo necesidad.
- **Reutilizando los patrones de diseño de ontologías (ODPs).-** Para este escenario se busca patrones existentes de diseño de ontologías y los reutiliza, es decir se diseña la ontología pero siguiendo patrones existentes.
- **Reestructuración de recursos ontológicos.-** En este escenario la ontología solo será reestructurada, es decir una ontología existente se extenderá, podará o especializará de acuerdo a la necesidad.
- **Localización de recursos ontológicos.-** Este escenario está basado en la modificación de una ontología basándose en el idioma y cultura de

la necesidad de la ontología, es decir se buscará una ontología multilingüe y que maneje términos similares.

Para cada uno de los escenarios antes mencionados, la metodología NeOn maneja los pasos de construcción de la ontología basada en el escenario seleccionado.

2.3.3.7. Análisis comparativo de las metodologías

Dentro de esta sección se realizará la selección de la metodología más apropiada para la construcción de la ontología que modelará el perfil de un usuario de Tv digital. Las diferentes metodologías descritas previamente, serán evaluadas numéricamente. El rango numérico que se utilizará serán valores enteros del 1 al 3, los cuales tendrán el siguiente significado:

- El valor numérico “1” representará el valor de “BAJO” cuando la metodología cumpla con un criterio de evaluación de manera básica.
- El valor numérico “2” representará el valor de “MEDIO” cuando la metodología cumpla con un criterio de evaluación a la mitad.
- El valor numérico “3” representará el valor de “ALTO” cuando la metodología cumpla con un criterio de evaluación de manera muy completa.

Los criterios de evaluación a los cuales serán sometidas cada una de las metodologías seleccionadas para esta evaluación, serán las siguientes:

- **Facilidad de Uso (FU).**- En este parámetro se evaluará la sencillez en los pasos para ser aplicados dentro de la metodología, es decir que los pasos estén bien definidos para ser realizados.
- **Polifuncionalidad por escenarios (PPE).**- La evaluación se realizará de acuerdo a la capacidad de la metodología para adaptarse a diferentes escenarios.
- **Documentación (DOC).**- Cantidad de documentación disponible acerca de la metodología a ser evaluada.
- **Reutilización (REU).**- Una vez aplicada la metodología facilidad de ser reusados sus pasos para la construcción de otro modelo ontológico.
- **Separación en los pasos (SP).**- Pasos bien definidos que no incluyan muchas tareas dentro de un mismo paso.
- **Guías Metodológicas Claras (GM).**- Que el contenido de las guías para usar la metodología para el usuario sean claras y consistentes.
- **Optimización de sus pasos (OP).**- Cada paso de la metodología esta optimizado de manera que cumpla más objetivos bien definidos dentro del mismo paso.



UNIVERSIDAD DE CUENCA

- **Documentación Actualizada (DA).**- Que la metodología haya sido actualizada de acuerdo a las nuevas necesidades.

Una vez definidos los parámetros y las metodologías a evaluar, en la tabla 2 se presenta los resultados de evaluación.

METODOLOGÍA	FU	PPE	DOC	REU	SP	GM	OP	DA	TOTAL
CYC	1	1	3	3	1	1	3	1	14
USCHOLD Y KING	3	2	3	1	3	1	2	1	16
GRÜNINGER Y FOX	3	1	3	1	2	2	3	2	17
KACTUS	2	1	3	2	1	2	2	2	15
METHONTOLOGY	2	2	3	2	3	2	2	2	18
NeOn	3	3	3	2	3	3	3	3	23

Tabla 2 Análisis Comparativo de Metodologías

Como se puede observar los resultados de la evaluación, la metodología NeOn es la que mayor puntuación obtiene por lo que será la utilizada en el desarrollo de la ontología para perfil de usuario en el capítulo 3 de esta tesis.



UNIVERSIDAD DE CUENCA

CREACIÓN DE LA ONTOLOGÍA DE PERFIL DE USUARIO

En el capítulo tres, se detallará el proceso de construcción de la ontología, siguiendo la metodología NeOn.

CAPÍTULO 3

CREACIÓN DE LA ONTOLOGÍA DE PERFIL DE USUARIO

En este capítulo se presentará paso a paso las diferentes actividades utilizadas para la creación de la ontología que será usada para el almacenamiento y consulta de un perfil de usuario para televisión digital. De acuerdo con el capítulo 2 se ha seleccionado la metodología NeOn para el proceso de creación de la ontología. El objetivo final del proceso metodológico es obtener una ontología que modele el perfil base de un usuario con los datos personales y sus gustos televisivos.

Dentro del capítulo se presentará inicialmente una introducción describiendo los objetivos planteados en el capítulo. Luego se especificarán los requerimientos de la ontología. Entonces, se describirá el proceso usado para modelar el perfil de usuario. El resultado del proceso metodológico será descrito como punto final del capítulo.

3.1. INTRODUCCIÓN

La necesidad de una recuperación de información más eficaz ha dado lugar a la creación de nuevas áreas de investigación como la Web semántica y la gestión de información personalizada. Estas áreas de estudio aprovechan el contexto semántico de los datos para facilitar el proceso de recuperación de la información. En este contexto, los perfiles de usuario juegan un rol vital para identificar las preferencias del usuario y personalizar las aplicaciones.

De forma general el proceso de personalización y contextualización de las preferencias de un usuario puede ser dividido en tres aspectos: i) la identificación de los datos a ser modelados y su formato de representación, ii) la determinación de las fuentes de información más apropiadas para la extracción de las preferencias y iii) el uso adecuado de este conocimiento para filtrar la información más relevante para el usuario.

El alcance de este capítulo está orientado hacia la búsqueda de una solución para el primero de estos tres aspectos. Es necesario hacer notar que la modelación del perfil del usuario planteada en este capítulo, tiene como meta

final proveer una base de conocimiento para un sistema de recomendación de programación televisiva basado en ontologías.

El enfoque ontológico para modelar el perfil de un usuario ha demostrado ser eficaz en la solución del problema del arranque en frío en los sistemas de recomendación, ya que permite la propagación de un pequeño número de conceptos iniciales a otros conceptos relacionados explotando la estructura ontológica del dominio [38], [39], [40], [41].

Por tanto, en este capítulo se explorará una representación semántica de las preferencias del usuario. Dicho esto, el modelamiento del perfil de usuario basado en ontologías plantea dos desafíos principales: la necesidad de conocimiento predefinido específico del dominio, por lo general construido manualmente, y correspondencias semánticas significativas entre la descripción del contenido y la información semántica en la ontología. En este capítulo se abordarán estos desafíos haciendo uso de una metodología que permita reutilizar, fusionar y ejecutar procesos de reingeniería de recursos ontológicos existentes para facilitar el proceso de creación de la ontología.

3.2. ESPECIFICACIÓN DE REQUERIMIENTOS DE LA ONTOLOGÍA

En esta sección se instancia la guía propuesta en la metodología NeOn para cubrir la actividad de especificación de requisitos de una ontología. El propósito es determinar los requisitos de la ontología que permita modelar las características y preferencias de un usuario de TV digital. Todas las actividades metodológicas se describen utilizando una plantilla y un flujo de trabajo. Tal como muestra la tabla 3, *para la plantilla* se debe especificar la definición de la actividad, junto con las características importantes como el objetivo, las entradas-salidas, etc. El uso de una plantilla ayuda a generar actividades de una manera práctica y fácil.

Especificación de requerimientos de la Ontología	
Definición La Especificación de requerimientos de la Ontología se refiere a la actividad de recolectar los requisitos que debe cumplir la ontología, por ejemplo, razones para construir ontología, grupo de objetivos, usos previstos, posiblemente alcanzado a través de un proceso de consenso.	
Objetivo Esta actividad indica por qué se está construyendo la ontología, cuáles son los usos que se tienen previstos para la ontología, quiénes son los usuarios finales, y cuáles son los requisitos que la ontología debe cumplir.	
Entrada El conjunto de las necesidades	Salida Documento de Especificación de



UNIVERSIDAD DE CUENCA

ontológicas	Requisitos de la Ontología (DERO)
Actores Desarrolladores de software y profesionales, que forman el equipo de desarrollo de ontologías (EDO), en colaboración con los usuarios y los expertos en el dominio.	
Cuando Esta actividad debe ser llevada a cabo al principio del proyecto y en paralelo con la actividad de adquisición de conocimiento.	

Tabla 3 Plantilla de Especificación de requerimientos de la ontología

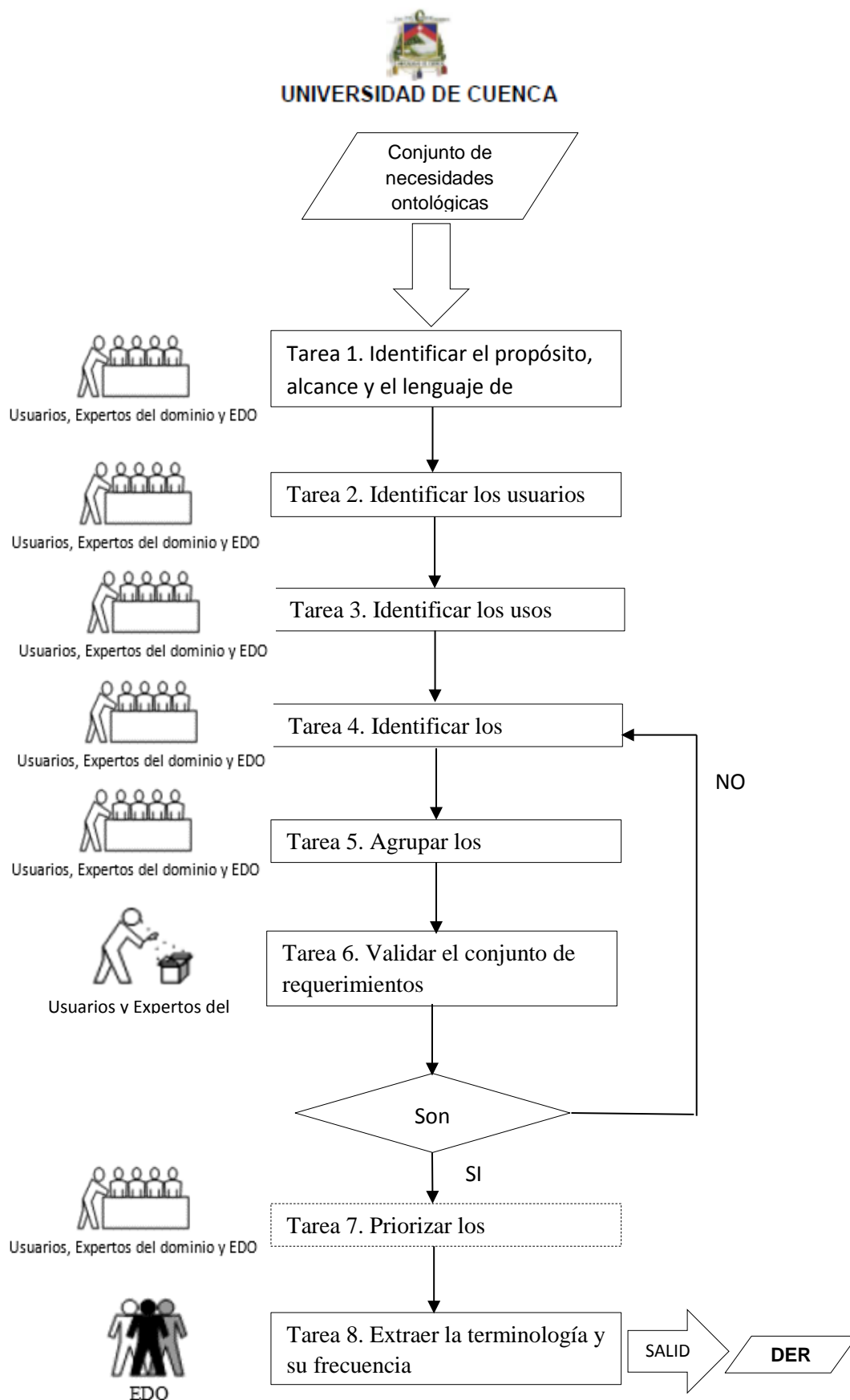


Figura 4 Tareas para requisitos Ontología Especificación

El flujo de trabajo es mostrado en la figura **4, la cual** representa los lineamientos metodológicos para la realización de la actividad de especificación de requisitos de la ontología de una manera detallada y prescriptiva. Esta, muestra las tareas principales involucradas, sus entradas, salidas y actores.

A continuación se irá describiendo los resultados de cada una de las tareas de la actividad de elicitación de requisitos, usando para ello la plantilla propuesta en la metodología. La Tabla **4** muestra las cinco primeras secciones del documento para el caso de la ontología. El propósito, alcance, y lenguaje de implementación de la ontología fueron determinados luego de una lluvia de ideas con los miembros del equipo de trabajo. Para determinar los posibles usuarios y los usos previstos de la ontología a desarrollar (tabla **3**), se tomó como base los casos de uso identificados para el desarrollo del sistema de recomendación que se está planeando implementar [42].

1	Propósito
El propósito de la construcción de la ontología del perfil de usuario es proveer un modelo consensuado del dominio de los usuarios de TV digital que pueda ser usado en un sistema de recomendación semántico.	
2	Alcance
La ontología tiene como alcance determinar las características de los de los usuarios de TV digital y sus preferencias televisivas tomando en consideración el contexto espacio-temporal en el que se ejecutan las actividades.	
3	Lenguaje de implementación
La ontología será implementada en OWL [43] y almacenada en un repositorio abierto.	
4	Usuarios Finales previstos
Usuario 1. Televidente que está buscando un programa para observar Usuario 2. Agencias televisivas que desean conocer las preferencias de los televidentes.	
5	Usos Previstos
Uso 1. Determinar las preferencias televisivas individuales tomando consideración el contexto (ej., día, género, localidad). Uso 2. Determinar las preferencias televisivas colectivas	

Tabla 4 Documento de especificación de requisitos: Secciones 1 a 5

En la identificación, agrupación y validación del conjunto de requisitos que debe satisfacer la ontología, participaron únicamente futuros usuarios del sistema de recomendación y no expertos del dominio. Para reducir el riesgo de una identificación errónea e incompleta de los requisitos, los resultados de las entrevistas fueron validadas por expertos del grupo de trabajo, todos ellos con sólidos conocimientos de ingeniería ontológica. Para identificar los

requerimientos funcionales se usó la técnica de preguntas de competencia [44] que consiste en escribir preguntas en lenguaje natural que la ontología a ser construida debe ser capaz de responder. La Tabla 5 muestra una lista ordenada de los requisitos funcionales y no funcionales.

6	Requerimientos de la ontología
a	No Funcionales
	<p>RNF1. La ontología debe soportar un escenario multilingüe en los siguientes lenguajes: inglés y español.</p> <p>RNF2. La ontología debe estar basada en estándares o modelos reconocidos en existencia o en desarrollo.</p>
b	Funcionales
	<p>Grupo 1 (Televidente)</p> <p>PC1. ¿Cuál es el nombre del televidente? Mauricio Espinoza</p> <p>PC2. ¿Cuál es la ciudad de residencia del televidente? Cuenca; Loja; Quito; Guayaquil, ...</p> <p>PC3. ¿Cuál es el rango de edad del televidente? niño; joven; adulto; adulto mayor</p> <p>PC4. ¿Cuáles son los tópicos de interés del televidente? deportes; noticias; series de acción</p> <p>PC5. ¿Qué ve el televidente en las mañanas-tardes-noches? Noticias nada-deportes; series de acción</p> <p>PC6. ¿Cuáles son sus programas favoritos? 24 Horas; FoxSport: Copa Libertadores; ESPN: Goles de Europa; CSI;</p> <p>PC7. ¿Qué tiempo dedica a ver la televisión en las mañanas –tardes-noches?</p> <p>30min-0min-60min</p> <p>PC8. ¿Qué ve el televidente los fines de semana? Futbol de Europa</p> <p>PC9. ¿Qué ve el televidente los feriados y días no laborables? Nada</p> <p>PC10. ¿Cuál es su personaje favorito? Lionel Messi</p> <p>PC11. ¿Cuál es el estado civil del televidente? casado</p> <p>PC12. ¿Cuál es la profesión del televidente? Ingeniero de Sistemas</p> <p>PC13. ¿Cuáles son las habilidades del televidente? Habla Inglés</p> <p>Grupo 2 (Grupo de Usuarios)</p> <p>PC14. ¿Qué ciudad ve más televisión? Guayaquil</p> <p>PC15. ¿Qué rango de edad ve más televisión? jóvenes</p> <p>PC16. ¿Qué tópicos de interés tienen los jóvenes-adultos-adultos? ¿Mayores? Películas -noticias; deportes-novelas</p> <p>PC17. ¿Qué ven los televidentes de la sierra-costa-oriental? noticias;</p>



UNIVERSIDAD DE CUENCA

deportes;-telenovelas; deportes- deportes- películas
PC18. ¿Qué ven los televidentes en la mañana -tarde-noche? noticias; Farándula –telenovelas – películas ;noticias; películas
PC19. ¿Cuál es el programa favorito de la Sierra? 24 horas
PC20. ¿Cuál es el programa favorito de la Costa? Televistazo
PC21. ¿Qué tiempo en promedio dedica a ver televisión la gente de Cuenca? 5 horas
PC22. Qué ve la gente de Cuenca los fines de semana? deportes
PC23. ¿Qué ve la gente de Cuenca los feriados? Películas; deportes

Tabla 5 Documento de especificación de requisitos: Sección 6

La última sección del documento de especificación de requisitos contiene el pre-glosario de términos dividido en tres partes: términos extraídos de las preguntas de competencia, términos provenientes de las respuestas a las preguntas de competencia, y términos identificados como nombres de entidades (objetos). Para la extracción del glosario de términos se adoptaron sencillas técnicas heurísticas de extracción de terminología.

Así por ejemplo, desde los diferentes requisitos se extrajeron nombres, adjetivos y verbos que serán representados en la ontología como conceptos, atributos, relaciones o instancias (en el caso de nombres de entidades). La Tabla 6 muestra las tres partes del glosario de términos. En las secciones *a* y *b* de la tabla, los números representan la frecuencia de aparición de los términos.

7	Pro-Glosario de Términos			
a	Términos desde las preguntas de competencia			
	Televidente	12	Ver	8
	Nombre	1	Ciudad	2
	Edad	2	Tópicos	2
	Programa	2	Mañana	3
	Noche	3	Tiempo	2
	Tarde	3	Fines de semana	2
	Gente	2	Feriado	2
	Estado civil	1	Personaje	2



UNIVERSIDAD DE CUENCA

	Profesión	1	Habilidad	1
	Costa	2	Sierra	2
	Cuenca	2	Oriente	2
b	Términos desde las respuestas a las preguntas de competencia			
	Cuenca, Loja, ...	1	joven, adulto, adulto mayor	1
	deportes, noticias, ...	4	30min, 60min, 90 min.	2
	soltero, casado, ...	1		
c	Objetos			
	Costa, Sierra, Oriente, etc... Cuenca, Loja, Quito, etc... Inglés, Español, etc.... Ingeniero de Sistemas, Arquitecto, etc.... 24 Horas, Televistazo, etc.... Lionel Messi, Neymar, etc....			

Tabla 6 Documento de especificación de requisitos: Sección 7

Aunque la metodología recomienda usar los términos con la frecuencia más alta para dirigir la búsqueda de recursos existentes que ayuden en el desarrollo de la ontología (ver sección 3.3-A), en este caso se procedió a categorizar la terminología con el fin de focalizar la búsqueda en dominios más específicos. Las categorías identificadas son: datos personales, preferencias, actividades, tiempo y lugares. A continuación se describen brevemente estas categorías.

Información del Televidente (Persona). Esta categoría contiene información acerca de la identidad del televidente. Esta categoría puede ser subdividida en:

- i) información básica del usuario como nombre, edad, sexo, estado civil, o lugar de residencia,
- ii) información profesional como sus títulos académicos,
- iii) Características propias del usuario como su peso o altura, y
- iv) Las habilidades del usuario como aficionado a la electrónica, manejo de idiomas, etc.

Tópicos de Preferencias. Esta categoría del perfil del usuario identifica las preferencias del usuario. Analizando las respuestas a las preguntas de

competencia, es posible identificar que un televidente puede tener al mismo tiempo diferentes preferencias (ej. deportes, noticias, etc.). Esta categoría deberá contener tópicos lo más generales posibles, de forma que sea factible almacenar cualquier tipo de preferencias.

Intereses. Algunos ejemplos de intereses que pueden ser útiles a la hora de recomendar programas televisivos son conocer que al usuario le gusta el rock, es aficionado a los videojuegos o que no le agradan los perros.

Actividades. La categoría actividades identifica toda la interacción que tiene el televidente frente al televisor. De hecho, los datos sobre el consumo de programas televisivos proporcionarán una gran cantidad de información acerca de las preferencias del usuario.

Tiempo y Lugares. Estas dos categorías son complementarias a las anteriores, en el sentido que permiten determinar el contexto espacio-temporal del usuario. Estas dos categorías permitirán entender mucho más al usuario, ofreciendo por ejemplo información precisa sobre el tiempo y lugar en los que un televidente observa un programa determinado.

Estas categorías de datos proveen la información necesaria para concentrar la búsqueda de modelos ontológicos que cubran estos dominios de interés.

3.3. MODELAMIENTO DEL PERFIL DEL USUARIO

En esta sección se explica el proceso seguido para crear un perfil semántico, basado en una ontología formal. Es necesario recordar que una ontología como definida en la sección 2.2, es una especificación formal y explícita de una conceptualización compartida. Aquí, *conceptualización* se refiere a un modelo abstracto de algún fenómeno en el mundo y la identificación de los conceptos relevantes de ese fenómeno. *Explícita* significa que el tipo de conceptos que se utilizan, así como las restricciones de su uso, se definen de forma explícita. *Formal* hace referencia al hecho de que la ontología debe ser legible por una máquina. *Compartida* refleja la noción de que una ontología captura conocimiento consensuado, es decir, que no es privado, sino aceptado por un grupo.

El uso de las ontologías para modelar el perfil de un usuario ha sido propuesto en varias aplicaciones como búsqueda en la Web [45], [46], gestión de información personal [47], o recomendación de noticias [48]. Inclusive, en el campo específico de este trabajo, existen varias propuestas que intentan modelar el perfil para un usuario de TV digital [49], [50]. A pesar de estos esfuerzos, no se puede hablar todavía de una ontología estándar para estos

propósitos, la mayoría de las ontologías creadas son específicas a la aplicación, cada una de ellas elaborada para un dominio particular.

El objetivo es reutilizar el trabajo existente en la literatura para crear una ontología que permita describir de manera adecuada el conocimiento relevante de un usuario de la TV digital. Se requiere que el modelo abarque tanto conceptualizaciones de las características del usuario (tales como profesión, edad, género, experiencia, etc.), como sus intereses televisivos, los mismos que pueden cubrir entidades de diferentes dominios.

A. Proceso de Construcción de la Ontología

Para este caso particular se ha adoptado el escenario seis de los nueve propuestos por la metodología NeOn (Sección 2.3.3.6): reusando, fusionando y ejecutando reingeniería de recursos ontológicos. Este escenario parte de la premisa que los desarrolladores tienen a su disposición recursos ontológicos útiles para modelar el dominio de su problema. Los recursos son fusionados para crear un nuevo modelo ontológico, sin embargo la fusión no es suficiente para cubrir las necesidades de la nueva ontología, por tanto un proceso de reingeniería es requerido.

A continuación se describen las actividades que se han llevado a cabo:

- 1) *Búsqueda de Ontologías*: Durante esta etapa se buscaron ontologías que reúnan los requisitos identificados en el *Documento de Especificación de Requerimientos de la Ontología*, introducido en la sección 3.2. Para la búsqueda de ontologías se utilizó Watson²³ y otros motores de búsqueda semántica como Swoogle²⁴ y Sindice²⁵. Se consultó en forma manual algunas librerías de ontologías existentes como Protégé Ontology Library²⁶. Adicionalmente, se buscaron estándares que definan tipos de competencia e intereses de un usuario en sitios Web de alta fiabilidad.

La tabla 7 muestra un ejemplo de las ontologías encontradas por cada categoría, incluyendo el proyecto o institución creador(a) del modelo. Para la categoría *información del televidente* se especificó la búsqueda en diferentes características que se requerían modelar acerca del usuario. El número entre paréntesis en cada categoría representa el total de ontologías encontradas y que están actualmente accesibles²⁷.

²³ <http://watson.kmi.open.ac.uk/>

²⁴ <http://swoogle.umbc.edu/>

²⁵ <http://sindice.com/>

²⁶ [http://protegewiki.stanford.edu/wiki/Protege Ontology Library](http://protegewiki.stanford.edu/wiki/Protege_Ontology_Library)

²⁷ El proceso de búsqueda inicio en el mes de octubre de 2013 y culmino a mediados de noviembre del mismo año.

2) *Comparación de las Ontologías.*: Durante esta tarea se compararon las ontologías identificadas en el paso anterior.

Para determinar el conjunto de ontologías no apropiadas a ser reutilizadas, las siguientes acciones fueron ejecutadas:

- Comprobar si el alcance y propósito establecido en el Documento de Especificación de Requisitos de la Ontología (DERO) es similar al de la ontología candidata.
- Comprobar si los requisitos no funcionales establecidos en el DERO están cubiertos por el dominio de la ontología candidata. Un ejemplo de requisito no funcional identificado en la propuesta fue que el multilingüismo debe ser representado en la ontología a desarrollar.
- Comprobar si los requisitos funcionales en forma de preguntas de competencia incluidos en el DERO son cubiertas (total o parcialmente) por el dominio de la ontología candidata. Esta comprobación fue ejecutada calculando la precisión y cobertura de la terminología de las ontologías candidatas con respecto a la terminología incluida en el DERO.

Se ha definido precisión como la proporción de los términos en la ontología candidata incluidos en los términos identificados en DERO sobre los términos en la ontología candidata. Esto se expresa de la siguiente manera:

$$Precision = \frac{TermOntCandidata \cap TermDERO}{TermOntCandidata}$$

La medida de cobertura está basada en la medida *recall* usada en recuperación de información. En este trabajo, la cobertura es la proporción de los términos identificados en el DERO que están incluidos en los términos recogidos en la ontología candidata sobre los términos identificados en el DERO. Esto se expresa de la siguiente manera:

$$Cobertura = \frac{TermOntCandidata \cap TermDERO}{TermDERO}$$

Con el fin de ilustrar el proceso ejecutado para analizar las ontologías candidatas bajo los criterios antes mencionados, se ha seleccionado como ejemplo el grupo de ontologías sobre información profesional del usuario (ver categoría 1 en la tabla 7). En la tabla se muestra el análisis efectuado



UNIVERSIDAD DE CUENCA

por el equipo de desarrollo de ontologías al grupo de recursos ontológicos seleccionado.

Ontología	Proyecto o Institución
Categoría 1: Información del Televidente (Persona)	
Información Básica (30)	
FOAF.owl GUMO.owl Person ²⁸	FOAF-Project DFKI GmbH W3C Schema.org
Información Profesional (7)	
FOAF.owl eBiquity Person ²⁹ Education Ontology ³⁰ eFOAF	FOAF-Project UMBC ebiquity Ontology Engineering Group
Características propias del usuario (2)	
GUMO LUMO	DFKI GmbH CERTH-ITI
Habilidades (12)	
GUMO Reusable Competency Definitions ³¹ Reusable Definition of Competency ³² HR-XML Person Competency ³³ Skill Ontology ³⁴	DFKI GmbH IEEE IMS HR-XML Consortium Ontology Engineering Group
Categoría 2: Tópicos de Preferencias (6)	
IPTC news codes ³⁵ DBPedia ³⁶	International Press Telecommunications Council Comunidad Wikipedia y otros
Categoría 3: Intereses (10)	
NERD ³⁷	G. Rizzo Y R. Troncy

²⁸ <http://schema.org/Person>

²⁹ <http://ebiquity.umbc.edu/ontology/person.owl>

³⁰ <http://mayor2.dia.fi.upm.es/oeg-upm/files/hrmontology>

³¹ <http://www.cen-ltso.net/main.aspx?put=264>

³² <http://www.imsglobal.org/competencies/>

³³ <http://www.hr-xml.org/>

³⁴ <http://mayor2.dia.fi.upm.es/oeg-upm/files/hrmontology>

³⁵ <http://webtlab.it.uc3m.es/results/NEWS/subjectcodes.owl>

³⁶ <http://wiki.dbpedia.org/Ontology>

Román Jarrín Sebastián Rodrigo

Sáenz Peñafiel Juan José

Interests ³⁸ E-FOAF ³⁹	M. Golemati
Categoría 4: Actividades (3)	
Event ⁴⁰ Programmes ⁴¹ Event ⁴²	Schema.org BBC Centre for Digital Music
Categoría 5: Tiempo (10) y Lugares (15)	
TimeLine Ontology ⁴³ Kestrel Time Ontology ⁴⁴ SRI Time Ontology ⁴⁵ OWL-Timer ⁴⁶ (antes DAML-Time)	Centre for Digital Music Kestrel Institute SRI Artificial Intelligence Center W3C
ISO3166 ⁴⁷ Geography Ontology ⁴⁸	ISO Ontology Engineering Group

Tabla 7 Ejemplo de ontologías por cada Categoría

Para decidir si una ontología candidata no es considerada útil en el proceso de reutilización, la siguiente heurística fue aplicada: Si el desarrollador contestó No a los criterios “Alcance similar” y/o “Propósito similar” y/o “Requerimientos Funcionales cubiertos”, entonces la ontología fue considerada no útil, y por tanto fue eliminada del conjunto de ontologías candidatas.

Criterio	Rango de Valores	Ontologías Candidatas							
		FO AF	eBiquity Person	Educación	Ukek ⁴⁹	Bibliographies ⁵⁰	Onto ⁵¹	KA ⁵²	

³⁷ <http://nerd.eurecom.fr/ontology>

³⁸ <http://oceanis.mm.di.uoa.g>

³⁹ <http://wiki.larkc.eu/e-foaf:interest>

⁴⁰ <http://schema.org/Event>

⁴¹ <http://www.bbc.co.uk/ontologies/programmes/2009-09-07.n3>

⁴² <http://motools.sourceforge.net/event/event.122.n3>

⁴³ <http://motools.sourceforge.net/timeline/timeline.179.n3>

⁴⁴ <http://www.kestrel.edu/DAML/2000/12/TIME.daml>

⁴⁵ <http://www.ai.sri.com/daml/ontologies/sri-basic/1-0/Time.daml>

⁴⁶ <http://www.w3.org/TR/owl-time/>

⁴⁷ http://www.iso.org/iso/country_codes

⁴⁸ <http://mayor2.dia.fi.upm.es/oeg-upm/files/hrmontology>

⁴⁹ <http://www.ukoln.ac.uk/projects/iemsr/wp2/rdf/vocab/uhec.rdf>

⁵⁰ <http://simile.mit.edu/repository/ontologies/official/bibliographies.rdfs>

⁵¹ <http://oaei.ontologymatching.org/2004/Contest/303/onto.rdf>

⁵² <http://www.cs.man.ac.uk/~horrocks/OWL/Ontologies/ka.owl>



Alcance similar	[Si, No, Desconocido]	SI	SI	SI	SI	NO	NO	NO
Propósito similar	[Si, No, Desconocido]	SI	NO	SI	NO	NO	NO	NO
Requerimientos no-funcionales cubiertos	[Si-Totalmente, Si-Parcialmente, No, NO, Desconocido]	NO	NO	NO	NO	NO	NO	NO
Requerimientos, funcionales cubiertos	[Si-Totalmente, Si-Parcialmente, No, Desconocido]	Si-Parcialmente	NO	Si-Parcialmente	NO	NO	NO	NO

Tabla 8 Ejemplo de comparación de ontologías de información personal

Luego de la aplicación de la heurística las ontologías: *eBiquity Person*, *Ukek*, *Bibliographies*, *Onto*, y *KA*, fueron eliminadas del conjunto de ontologías candidatas. El mismo proceso fue aplicado para el resto de ontologías identificadas en las otras categorías.

3) *Selección de las Ontologías*: En esta tarea se determinó cuál de las ontologías identificadas en la tarea previa es la más apropiada para ser reutilizada en la ontología a ser desarrollada. Para determinar tal ontología se analizaron cuatro dimensiones de las características no-funcionales de las ontologías. Las dimensiones consideradas se describen brevemente a continuación:

- Costo de Reutilización. Se refiere a la estimación del coste (económico y temporal) que se necesita para la reutilización de la ontología candidata.
- Esfuerzo de Comprensión. Se refiere a la estimación del esfuerzo necesario para comprender la ontología candidata.
- Esfuerzo de Integración. Se refiere a la estimación del esfuerzo necesario para la integración de la ontología candidata en la ontología que está siendo desarrollada.
- Confiabilidad. Se refiere al análisis de si se puede confiar en la ontología candidata a ser reutilizada.

Es necesario aclarar que cada una de las dimensiones arriba descritas posee un conjunto de criterios que permiten obtener un puntaje para cada ontología candidata. Por cada criterio se tiene i) un rango de valores lingüísticos⁵³, ii) una explicación de la forma de medir el criterio, y iii) un peso numérico. Una explicación de los criterios y su forma de medición esta fuera del alcance de este artículo, sin embargo mayores detalles pueden ser consultados en [51].

Criterio	Peso	Valores54	
		FOAF	Education
Costo de Reutilización			
Costo Económico del Reuso	(-) 9	1	1
Tiempo Requerido para Reuso	(-) 7	1	1
Esfuerzo de Comprensión			
Calidad Documentación	(+) 8	3	2
Disponibilidad de Conocimiento Externo	(+) 7	3	2
Claridad de Código	(+) 8	3	3
Esfuerzo de Integración			
Adecuación de extracción de conocimiento	(+) 9	2	2
Adecuación de convenciones de nombrado	(+) 5	2	3
Adecuación del lenguaje de implementación	(+) 8	3	3
Conflictos de conocimiento	(-) 8	2	2
Adaptación de razonador	(+) 5	3	1
Necesidad de términos puente	(-) 7	1	1
Confiabilidad			
Disponibilidad de pruebas	(+) 8	3	1
Soporte teórico	(+) 8	3	3
Reputación del equipo de desarrollo	(+) 8	3	3
Fiabilidad de documentación	(+) 3	3	3
Apoyo Práctico	(+) 6	3	3
	Total	1.48	1.26

Tabla 9 Ejemplo de selección de ontologías de información personal.

La tabla 9 muestra los valores obtenidos para cada criterio y por cada ontología candidata de la categoría información personal. Considere que este mismo

⁵³ Estos valores lingüísticos toman valor en una escala de [0 a 3] de la siguiente manera: desconocido=0, bajo=1, medio=2, y alto=3

⁵⁴ desconocido=0, bajo=1, medio=2, y alto=3

grupo de ontologías fue usado como ejemplo para mostrar la tarea de comparación de ontologías (descrito en el punto anterior). Del grupo de siete ontologías, sólo dos fueron catalogadas como aceptables para ser reusadas. En la tabla 9, los pesos numéricos fueron definidos por el equipo de trabajo, considerando la importancia de cada uno de los criterios. Los símbolos (+) y (-) en los pesos se establecieron para observar si el criterio se cuenta de manera positiva o negativa, respectivamente. Tanto los símbolos como los pesos numéricos no fueron modificados por el desarrollador de la ontología al momento de la evaluación. Para calcular la puntuación de las diferentes ontologías candidatas se usaron las siguientes formulas:

$$Puntaje_{i(+)} = \sum_{j(+)} Valor_{Ti,j} x \frac{Peso_j}{\sum_j Peso_j}$$

$$Puntaje_{i(-)} = \sum_{j(-)} Valor_{Ti,j} x \frac{Peso_j}{\sum_j Peso_j}$$

Donde:

- *Puntaje (+)* es el puntaje para la ontología candidata *i* para el conjunto de criterios ponderados con (+)
- *Puntaje (-)* es el puntaje para la ontología candidata *i* para el conjunto de criterios ponderados con (-)
- *i* es una ontología candidata particular
- *j* es un criterio particular de las incluidas en la Tabla 9, *j (+)* representa los criterios con peso positivo y *j (-)* los criterios con peso negativo.
- *Valor_{Ti,j}* es el valor para el criterio de *j* en la ontología *i*
- *Peso_j* es el peso numérico asociado al criterio *j*

La puntuación final para cada ontología candidata se obtuvo de la siguiente manera:

$$Puntaje_i = Puntaje_{i(+)} - Puntaje_{i(-)}$$

Las ontologías con el puntaje más alto en cada una de las categorías fueron seleccionadas para la construcción del perfil de usuario. En ciertos casos más de una ontología fue seleccionada dada su calidad y aporte al proceso de construcción.

- 4) *Personalización e Integración de las Ontologías Seleccionadas*: Para personalizar e integrar las ontologías seleccionadas se ejecutó un proceso de alineamiento manual. El objetivo fue crear un modelo ligero y manejable que i) pueda soportar la representación del perfil de un usuario de TV digital bajo un único vocabulario, ii) abarque el conjunto mínimo de conceptos entre la abundante información en el dominio, y iii) sea capaz de mantener conceptualizaciones específicas tales como características y preferencias del usuario.

El proceso de alineamiento no tiene sentido sin un proceso de re-estructuración de los recursos ontológicos a ser alineados.

Algunas situaciones pueden darse:

- Los recursos importados contienen información irrelevante para modelar el perfil.
- Algunos conceptos relevantes para modelar el perfil no están contempladas en los recursos importados.

Con el fin de cumplir los principios antes descritos, algunas de estas sub-actividades fueron ejecutadas en los recursos ontológicos seleccionados para reutilización:

- *Podar las ontologías de acuerdo a las necesidades*. En algunos casos fue necesario eliminar ciertos conceptos que no aportaban nada a la ontología construida. Por ejemplo en el caso de la ontología GUMO todos los conceptos derivados a partir del concepto Personality fueron eliminados porque no aportaban nada a las necesidades de la ontología final.
- *Enriquecer la ontología a ser reusada*. En ciertas ontologías seleccionadas fue necesario ejecutar un proceso de enriquecimiento con nuevas estructuras conceptuales que permitan cubrir las necesidades puntuales de la ontología a ser desarrollada.
- *Adaptar las ontologías seleccionadas a los criterios de diseño de la ontología destino*. Algunas convenciones de nombrado de términos fueron aplicadas para homogenizar la ontología final.
- *Incluir conexiones a través de relaciones entre clases que no fueron inicialmente conectadas*. Por ejemplo, para conectar la ontología sobre la información básica de la persona con la ontología que describe la localidad de residencia del usuario se usó la propiedad *viveEn (liveIn)*.

En ningún caso fue necesario traducir una ontología a ser reusada al lenguaje de implementación seleccionado para la ontología final. Todas las ontologías estaban disponibles en formatos similares ya sea RDF u OWL.

La figura 5 muestra el metamodelo de la ontología que define el perfil de un usuario de TV digital. Los rectángulos en la figura representan clases, mientras que los óvalos representan relaciones. El metamodelo no es completo, solo pretende ilustrar las relaciones entre las diferentes ontologías usadas para modelar el perfil.

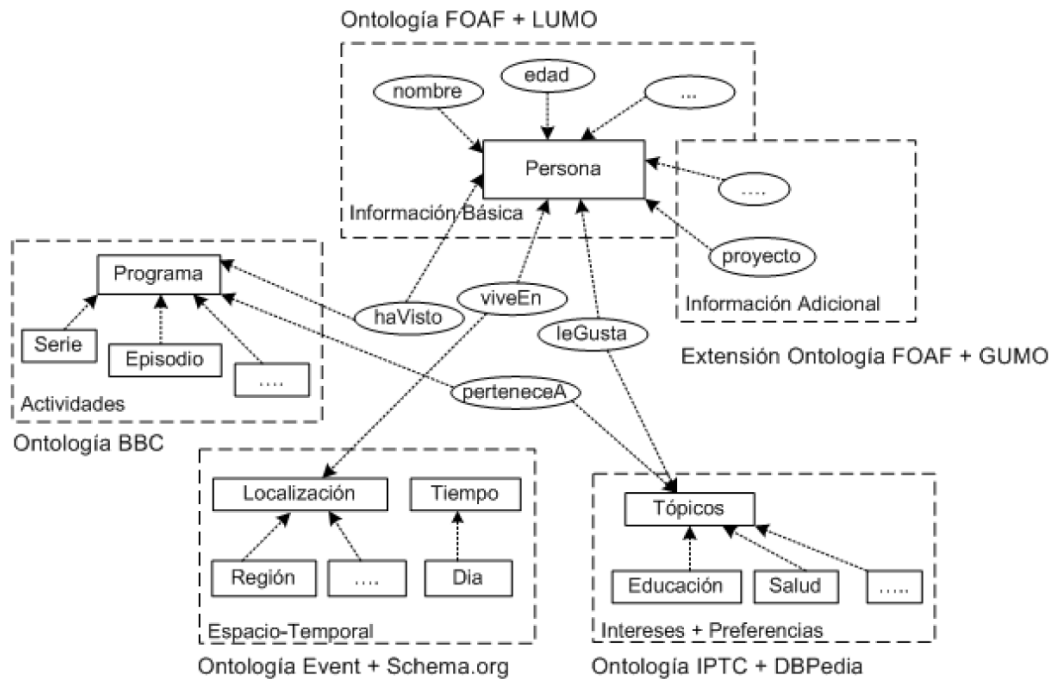


Figura 5 Metamodelo de la Ontología del Perfil de Usuario

La ontología actual tiene 585 Classes, 9 ObjectProperties y 30 DataProperties.

- 5) *Evaluación de las Ontología Resultante:* Esta actividad tiene como meta comprobar la calidad técnica de la ontología construida. Hay que indicar que en este trabajo únicamente se ha comprobado que la ontología no contenga errores de sintaxis.

En este capítulo se identificó los datos a ser modelados y su formato de representación, la determinación de las fuentes de información más apropiadas para la extracción de las preferencias del usuario será descrito en el siguiente capítulo.



UNIVERSIDAD DE CUENCA

RECOLECCIÓN DEL PERFIL DE USUARIO DESDE FUENTES EXTERNAS

En este capítulo, se detallará el método de recolección de datos externos que fue utilizado en el desarrollo de la tesis.

CAPÍTULO 4

RECOLECCIÓN DEL PERFIL DE USUARIO DESDE FUENTES EXTERNAS

En el presente capítulo se detalla la recolección de los datos del usuario, explicando los procedimientos y métodos seleccionados que permiten poblar el modelo ontológico del perfil de usuario para la televisión digital. Primero se presenta la metodología seleccionada para la recolección de datos (*Híbrida*), continuando con un breve repaso de algunos métodos de acceso *explícitos* e *implícitos* y la selección de los métodos de acceso a los datos que se utiliza en el proyecto. Finalizando con la explicación de *los FORMULARIOS WEB* y el desarrollo del módulo de redes sociales unificado *MRS*, junto con las ventajas más relevantes de cada uno de estos métodos.

4.1. METODOLOGIA

Como se ha visto en el capítulo 2 sección 2.1.2 existen 3 metodologías para la extracción de datos: implícita, explícita e híbrida. Para el alcance de este proyecto se ha decidido, debido a las ventajas presentadas en la sección **2.1.2.3**, utilizar el método **HÍBRIDO**. En la figura **6** se observa la organización del método híbrido, que nos permite recolectar información más precisa, pues se dispone de una fuente de datos que proviene de manera directa por el usuario (*METODO EXPLICITO*) y otra fuente alternativa de datos que proviene de manera indirecta desde el usuario (*METODO IMPLICITO*). Una ventaja del método *híbrido* es que el usuario puede tener alternativas para utilizar la técnica que desee para la entrega de sus datos.

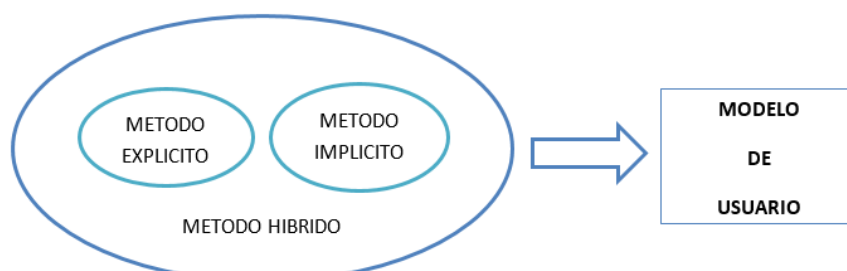


Figura 6 Técnica de recolección de información Híbrida

4.2. TÉCNICAS PARA RECOLECCIÓN DE DATOS

Existen diversas técnicas para recolectar información de un usuario específico, que van desde formularios y entrevistas, hasta recolección de datos de redes sociales y eventos de la televisión, a continuación se detallan las diferentes técnicas para recolectar información del usuario usando métodos explícitos e implícitos.

4.2.1. Información recolectada con técnicas explícitas.

Los datos explícitos se proporcionan de manera directa desde el usuario, es decir que la precisión y consistencia de la información que se recolecta depende del usuario. Existen varias metodologías explícitas, donde el usuario puede brindar información. Entre estos están:

Fuente de datos	Comentario
Formularios	Un formulario es un documento ya sea físico (en papel) o digital (web), diseñado con el propósito de que el usuario introduzca datos estructurados (nombre, apellidos, dirección, etc.) para luego ser almacenados y procesados.
Encuestas	Es un método basado en un cuestionario previamente diseñado, parecido a formulario pero más complejo, los datos se obtienen realizando un conjunto de preguntas normalizadas dirigidas a una muestra representativa o al conjunto de personas
Observación	Es un método general, se recolecta solo observando y analizando los datos de un elemento utilizando los sentidos como instrumentos principales.
Entrevistas	Es un método que se basa en un diálogo entablado entre dos o más personas.

Tabla 10 Técnicas de recolección explícita de datos

4.2.2. Información recolectada con técnicas implícitas.

La información implícita se obtiene de manera indirecta desde el usuario, o parcialmente indirecta ⁵⁵. En este tipo de información, la precisión y consistencia de datos depende de la fuente o el método de recolección. Existen varias técnicas implícitas, donde se puede obtener información del usuario. Entre estos están:

Fuente de datos	Comentario
Redes Sociales	Es un servicio online que es utilizado como medio de comunicación social, se centra en encontrar gente para relacionarse en línea. Formado por personas que comparten sus datos entre otros.
Eventos Televisión	Compuesto de la interacción del usuario con la televisión digital, este método se centra en captar la “interacción” que realiza la persona en la televisión, tales como el tiempo que mira un canal, el día o la hora, etc.
Ingeniería Social	Es un método que se centra en la práctica de obtener información confidencial a través de la manipulación de usuarios. Como ejemplo, un ingeniero social usa comúnmente el teléfono o Internet para engañar a la gente, fingiendo ser, por ejemplo, un empleado de algún banco o alguna otra empresa, un compañero de trabajo, un técnico o un cliente para obtener información de dicho usuario.

Tabla 11 Técnicas de recolección implícita de datos.

4.3. RECOLECCIÓN DE DATOS EXPLÍCITOS

Como se describió en la sección 4.2.1, existen diferentes técnicas para recolectar información del usuario de manera *explícita* como: formularios, encuestas, observación y entrevistas. Para el alcance de este proyecto, se

⁵⁵ El usuario puede permitir recolectar cierta información y no la totalidad de la misma.

decidió obtener información explícita del usuario desde un **FORMULARIO WEB**, debido a las ventajas que ofrece este técnica frente a los otros métodos. Algunas de estas ventajas son explicadas en la tabla 12.

Ventajas	Comentario
<i>Facilidad de uso para el usuario</i>	Los formularios web se encuentran en numerosas páginas en internet, el usuario está acostumbrado a utilizarlos y por ende, tiene más experiencia en utilizarlos.
<i>Obtener datos con mayor precisión</i>	Un formulario web puede hacer uso de combo box, list box, calendarios, etc. que facilitan y mejoran la precisión de los datos.
<i>Mejor recolección de datos</i>	Manejo directo de información, desde el formulario web al modelo de usuario, se evade perder/errar información.
<i>Respaldo de información</i>	También una ventaja de utilizar un formulario web, es que nos permite respaldar información
<i>Velocidad de uso</i>	La mayoría de navegadores web almacenan datos de formularios previamente llenados en la cache, esta información sirve para reutilizarlos en otros formularios, brindando facilidad y optimizando tiempos.

Tabla 12 Ventajas de un formulario Web.

En [52] el autor propone guías para el diseño de un formulario web, estas son pequeñas recomendaciones para poder crear un formulario web, permitiendo eficiencia a la hora de recolectar información. Un resumen de estas guías se presenta en la tabla 13.

Recomendaciones para el diseño de formularios WEB	
Para presentación de Información	<ul style="list-style-type: none"> • Para reducir los tiempos de llenar el formulario, y cuando son datos conocidos utilizar: <i>top aligned</i>. • Cuando el espacio vertical de la pantalla es una restricción utilizar: <i>right aligned</i>. • Para datos o conocidos o datos complejos utilizar: <i>left</i>



UNIVERSIDAD DE CUENCA

	<i>aligned.</i>
Para campos obligatorios en el formulario	<ul style="list-style-type: none">• Evitar siempre campos <i>opcionales</i>.• Si la mayoría de campos son requeridos: indicar campos opcionales.• Si la mayoría de campos son opcionales: indicar campos requeridos.• El texto es lo mejor, pero el signo “*” sirve para indicar los campos requeridos• Utilizar indicadores asociados con etiquetas.
Para el ancho de campos	<ul style="list-style-type: none">• Cuando sea posible, usar el ancho del campo como <i>affordance</i>⁵⁶• De lo contrario considerar una longitud coherente que proporcione suficiente espacio para digitar los datos.
Para el Contenido Agrupado	<ul style="list-style-type: none">• Usar agrupaciones de contenido relevante para organizar los formularios.• Usar el mínimo de elementos visuales necesarios para comunicar relaciones disponibles.
Para las Acciones	<ul style="list-style-type: none">• Evitar acciones secundarios de ser posible• De lo contrario, asegurar una clara distinción entre las acciones primarias y secundarias del formulario• Alinear las acciones primarias con campos que den una idea clara de la acción.
Para la Ayuda y consejos en los formularios	<ul style="list-style-type: none">• Minimizar la cantidad de ayuda y consejos requeridos para llenar el formulario• La ayuda visible y adyacente a los datos solicitados es más útil.• Cuando se solicita gran cantidad de datos desconocidos, considerar mejor un sistema de ayuda dinámico.
Para la Interacción	<ul style="list-style-type: none">• Remover todas las solicitudes de datos innecesarias• Activar valores por defecto• Emplear entradas de datos flexibles• Aclarar un camino fácil para completar el formulario.• Para formularios largos: mostrar el progreso y el guardado
Para la Tabulación	<ul style="list-style-type: none">• Recordar tener en cuenta el comportamiento de la tabulación• Usa el atributo <i>tabindex</i> para controlar el orden de tabulación.• Considerar expectativas de tabulación cuando se diseñen los formularios.
Para la Activación de campos en tiempo real	<ul style="list-style-type: none">• Mapear la activación de campos para priorizar las necesidades del usuario• Más efectivo cuando el usuario lo inicia.• Mantener un enfoque coherente.

⁵⁶ Affordance es la cualidad de un objeto o ambiente que permite a un individuo realizar una acción.

Recomendaciones Adicionales	<p>Validaciones en línea</p> <ul style="list-style-type: none"> • Utilizar validación en línea para campos que tengan un alto riesgo de errores • Sugiera entradas de campos para evitar ambigüedad • Explicar los límites de los campos <p>Errores</p> <ul style="list-style-type: none"> • Comunicar claramente cuando un error ha sucedido: mostrar en la parte superior, resaltado el error. • Proveer acciones para corregir los errores. • Asociar los campos responsables con el error principal. <p>Progreso</p> <ul style="list-style-type: none"> • Proveer indicadores de las tareas en ejecución. • Deshabilitar el botón “submit” después que el usuario haya hecho clic para evitar duplicar el procesamiento <p>Finalización</p> <ul style="list-style-type: none"> • Comunicar claramente que los datos fueron procesados • Proveer información de acuerdo al contexto de los datos procesados

Tabla 13 Guías para el diseño de un formulario Web.

Un gran parte de estas recomendaciones han sido consideradas en la elaboración del formulario de datos personales descrito en la siguiente sección.

4.3.1. Formulario Web, Datos personales

Después de seleccionar la metodología de recolección de datos explícita, el diseño del **FORMULARIO WEB** se ajustará a los datos personales del modelo de usuario, especificado en el **capítulo 3**, dicho modelo se detalla a continuación en la tabla 14.

Atributo	Tipo	Tamaño
Nombre	Cadena e texto	50
E-mail	Cadena de texto	50
Edad	Numero	11
Ubicación	Cadena de texto	50
Ocupación	Cadena de texto	50
Sexo	Cadena de texto	10

Tabla 14 Atributos del perfil de usuario

En base a las recomendaciones descritas por Wroblewski [52] se realizó el diseño principal del formulario web para los datos personales del usuario, En la figura 7 se detalla el bosquejo inicial del diseño. Para mayor precisión de datos del usuario, el formulario web debe contener el menor número de preguntas abiertas⁵⁷, para evitar datos inconsistentes y ruido en la información.

4.3.2. Formulario Web, Preferencias

Así como los datos personales, el diseño del **FORMULARIO WEB** se ajustará a las preferencias del modelo de usuario, especificado en el capítulo 3, dicho modelo se detalla a continuación en la tabla 15.

Atributo	Tipo	Tamaño
Descripción	Cadena de texto	50
Categoría	Numero	11
Calificación	Cadena de texto	50

Tabla 15 Atributos de las preferencias del usuario

El formulario web para las preferencias se basa en las recomendaciones de la tabla 13 para la presentación de Información: para campos obligatorios en un formulario, recomendaciones para diseñar contenido agrupado, y recomendaciones adicionales. En la figura 8 se detalla el bosquejo inicial del diseño del formulario web. Para mayor precisión de datos del usuario, el formulario web también debe contener el menor número de preguntas abiertas.

⁵⁷ Una pregunta abierta le solicita al consultado que formule su propia respuesta sin delimitar la misma.

FORMULARIO

NOMBRE

E-MAIL

EDAD

UBICACION

OCUPACION

SEXO ☐ MASCULINO ☐ FEMENINO

Text box

Spinner

Combo box

Check box

Figura 7 Diseño del formulario web para los datos personales del usuario

FORMULARIO

DESCRIPCION

CATEGORIA

CALIFICACION

Text box

Combo box

Spinner

Figura 8 Diseño del formulario web para las preferencias del usuario

4.3.3. Esquema general de los formularios Web

En el diseño presentado en las secciones 4.3.1 y 4.3.2 se especificaron las características necesarias de los **FORMULARIOS WEB**. Basado en los argumentos de diseño y el modelo de datos personales e intereses de usuario (ver tablas **14 y 15**), se procedió a implementar el formulario web según el diseño representado en las figuras **7 y 8**.

Para poder crear un formulario web, se debe tener en cuenta: el *servidor web*⁵⁸ que responderá la solicitud de proceso de datos, el *lenguaje de programación* que procesará el servidor, y el método de acceso. La implementación de estos formularios se detallará en el *capítulo 5 Prototipo de la aplicación del perfil de usuario*.

4.4. RECOLECCIÓN DE DATOS IMPLÍCITOS

Como se analizó en la sección 4.2.2, existen muchas técnicas para recuperar información *implícita* de un usuario como: redes sociales, eventos en la televisión, ingeniería social, etc. Para el alcance de este proyecto, se ha decidido obtener la información *implícita* del usuario, a través de las **REDES SOCIALES** y **EVENTOS DE LA TELEVISIONDIGITAL**. *Un resumen de las principales* ventajas de estos métodos es presentado en las tablas **16 y 17**.

Ventajas Redes Sociales	Comentario
<i>Fuente de datos confiable</i>	Al ser una red social, los usuarios comparten información segura entre sus contactos.
<i>Facilidad de acceso</i>	La mayoría de redes sociales, tienen una interfaz de acceso que permite con mayor facilidad acceder a los datos de los usuarios
<i>Uso mayoritario</i>	Cada vez el uso de las redes sociales está siendo más extendido en el mundo, logrando en un punto tener un mayor alcance y cantidad e información.

⁵⁸ Un servidor web es un programa informático que procesa una solicitud del lado del servidor.
 Román Jarrín Sebastián Rodrigo
 Sáenz Peñafiel Juan José



UNIVERSIDAD DE CUENCA

<i>Diversidad de información</i>	Existe un sin número de tipos de red social que nos permiten obtener diferentes tipos de datos como: seguidores, gustos, actividades, etc.
----------------------------------	--

Tabla 16 Ventajas de las Redes Sociales

<i>Ventajas TDT⁵⁹</i>	<i>Comentario</i>
<i>Más variedad de contenido</i>	TV Digital permite programas, que se escogen de una guía Electrónica de Programación, con un menú sobre la programación disponible, haciendo que el usuario pueda escoger de mejor manera sus gustos.
<i>Usuarios evitan lo que no le interesa</i>	El usuario en la televisión digital evita los programas que no son de su agrado, estos datos son importantes al momento de filtrar los intereses de un perfil de usuario.
<i>Usuarios profundizan lo que le interesa</i>	Caso contrario al anterior, el usuario en la televisión digital mira durante más tiempo los programas que son de su agrado, ayudando así a la recolección de gustos y preferencias.

Tabla 17 Ventajas de los Eventos de la televisión Digital.

4.4.1. Redes sociales

En la actualidad, según el INEC⁶⁰ (Instituto Ecuatoriano de Estadísticas y Censo) en su última presentación de resultados (2013), en Ecuador el 28,3% de los hogares a nivel nacional tienen acceso a internet, mientras el 40,4% de la población de Ecuador utilizó Internet los últimos 12 meses. Otro dato interesante que presenta el INEC es que el 28.2% de los ecuatorianos que tienen acceso a internet, lo utiliza como medio de comunicación, es decir entre redes sociales, correo, etc. Quizás el dato más relevante para los objetivos del proyecto sea que el 6.81% de la población del Ecuador utiliza redes sociales, 2 puntos más que el 2012, y 4 más que el 2011, como se observa en la figura 9.

Es notorio el crecimiento de las Tecnologías de Información y Comunicación (TIC's) en Ecuador, sobre todo el acceso a internet y la utilización de las redes sociales. Cada año aproximadamente 300,000 ecuatorianos crean un nuevo

⁵⁹ Televisión Digital Terrestre

⁶⁰ <http://www.inec.gob.ec/home/>

Román Jarrín Sebastián Rodrigo

Sáenz Peñafiel Juan José

usuario en una red social, es por eso que las redes sociales es una fuente muy importante de datos.

Como se observó en la tabla 1 del capítulo 2 para el alcance del proyecto, fueron seleccionados 3 redes sociales principales como fuente de datos: **Facebook, Twitter y Google+**. A continuación se detalla los aspectos importantes de cada red social.

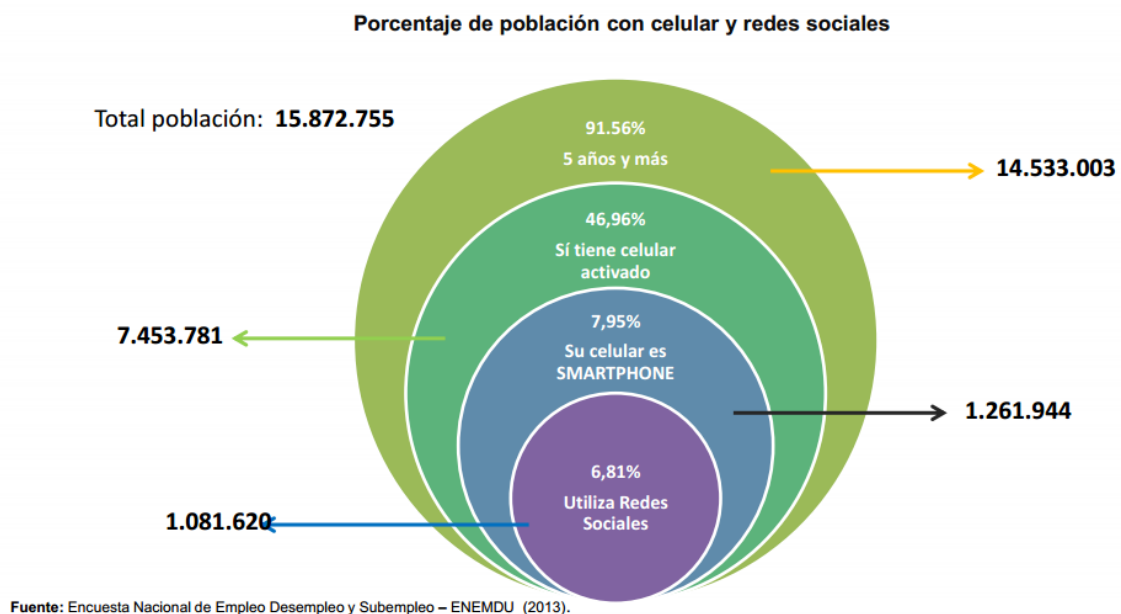


Figura 9 Porcentaje de población con celular y redes sociales según INEC

4.4.1.1. Especificaciones técnicas de las Redes sociales.

Después de definir las redes sociales como una de las metodologías para acceder a datos implícitos, se detallan las características técnicas, que permiten recuperar los datos del usuario. En la tabla 18 se definen los lenguajes que soporta la API⁶¹, así como características importantes como el protocolo de autenticación y si existe una buena documentación.

Red social	API	PHP	JS P	AS P	JavaScript	Protocolo	Documentación	Comentario
------------	-----	-----	------	------	------------	-----------	---------------	------------

⁶¹ API es el conjunto de funciones y procedimientos que ofrece cierta biblioteca, en este caso red social, para ser utilizado por otro software como una capa de abstracción
 Román Jarrín Sebastián Rodrigo
 Sáenz Peñafiel Juan José

Facebook	Si	Si	No	No	Si	OAuth 2.0	Si, Excelente	Buena documentación, Soporte PHP, JS
Twitter	Si	Si	No	No	Si	OAuth	Si, Excelente	Buena documentación, Soporte PHP, JS
Google+	Si	Si	Si	No	Si	OAuth 2.0	Si, Excelente	Buena documentación, Soporte PHP, JSP, JS

Tabla 18 Especificaciones técnicas de las Redes sociales.

4.4.1.2. Protocolos de comunicación

Para poder obtener datos de una red social, se necesita un **protocolo de comunicación** entre cualquier aplicación y la *API* de dicha red social. Como observamos en la tabla **18**, existen diversos protocolos de autorización que utilizan las *redes sociales* para el proceso de autenticación, a continuación detallamos los más relevantes.

- **Protocolo OAuth**

OAuth⁶² (Open Authorization) es un protocolo abierto, propuesto por Blaine Cook y Chris Messina, que permite autorización segura de una API de modo estándar y simple para aplicaciones de escritorio, móviles y web. OAuth proporciona a los usuarios un acceso a sus datos al mismo tiempo que protege las credenciales de su cuenta.

OAuth permite a un usuario del sitio A compartir su información en el sitio A (proveedor de servicio) con el sitio B (llamado consumidor) sin compartir toda su identidad. La figura **10** muestra el flujo abstracto del protocolo OAuth.

⁶² <http://es.wikipedia.org/wiki/OAuth>



UNIVERSIDAD DE CUENCA

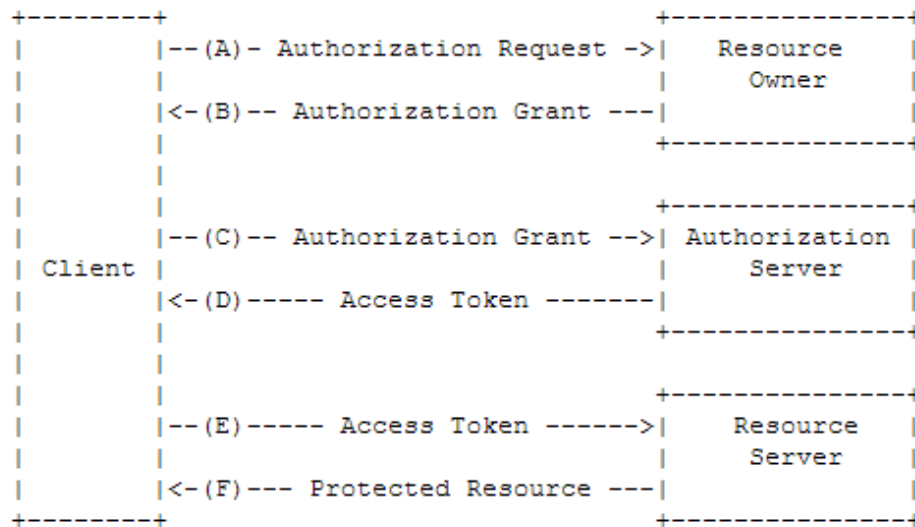


Figura 10 Flujo del protocolo OAuth ⁶³

- **Protocolo OpenID**

OpenID es un estándar de identificación digital descentralizado, con el que un usuario puede identificarse en una página web a través de una URL (o un XRI en la versión actual) y puede ser verificado por cualquier servidor que soporte el protocolo.

En los sitios que soporten OpenID, los usuarios no tienen que crearse una nueva cuenta de usuario para obtener acceso. En su lugar, solo necesitan disponer de un identificador creado en un servidor que verifique OpenID, llamado *proveedor de identidad* o *IdP*.

El proveedor de identidad puede confirmar la identificación OpenID del usuario a un sitio que soporte este sistema. La figura **11** muestra el flujo abstracto del protocolo OpenID.

⁶³ <http://tools.ietf.org/html/rfc6749>

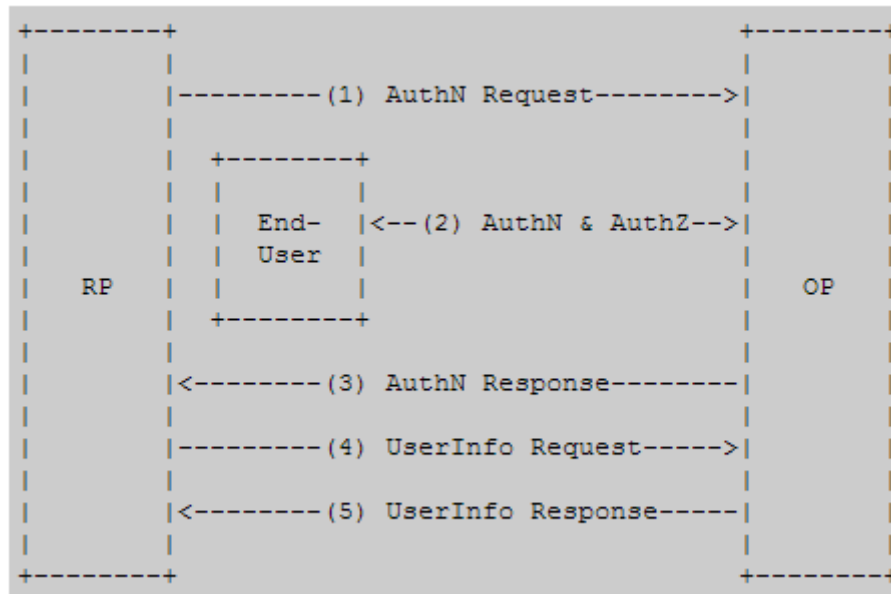


Figura 11 Flujo del protocolo OpenID⁶⁴ RP (Cliente), OP (Proveedor OpenID)

4.4.1.3. Procesos de autenticación

Después de ver los protocolos de autorización, otra parte importante del acceso a los datos del usuario en las redes sociales, es el proceso de **autenticación**, que se basa en un protocolo de autorización descrito en la sección anterior, a continuación detallamos cada uno de ellos.

- **Proceso de autenticación Facebook**

La red social **FACEBOOK** se basa en el protocolo **OAuth 2.0**, como vimos anteriormente este protocolo, genera una petición de autorización, el proveedor genera una solicitud de permisos, y según el usuario acepte o no, se genera una clave denominada **"access token"** que permite acceder a los datos del usuario. La figura **12** muestra el flujo de petición de un **access token** en Facebook.

⁶⁴ http://openid.net/specs/openid-connect-core-1_0.html

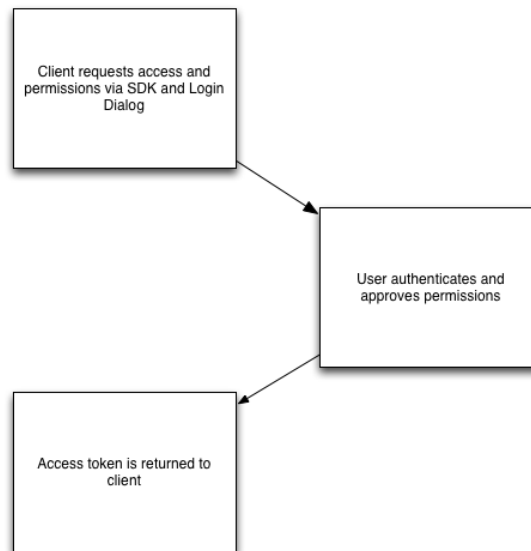


Figura 12 Flujo de solicitud de acceso de Facebook⁶⁵

- **Proceso de autenticación Twitter**

La red social **TWITTER** también se basa en el protocolo **OAuth 2.0**, y el proceso de autenticación es similar a *Facebook*, la aplicación verifica si está identificado el usuario y si tiene los permisos, caso contrario se genera una petición de autorización, el proveedor genera una solicitud de permisos a su vez, y según el usuario acepte o no, se genera el **access token** necesario para acceder a los datos del usuario. La figura **13** muestra el flujo de petición de un **access token** en Twitter.

⁶⁵ <https://developers.facebook.com/docs/facebook-login/access-tokens>



UNIVERSIDAD DE CUENCA

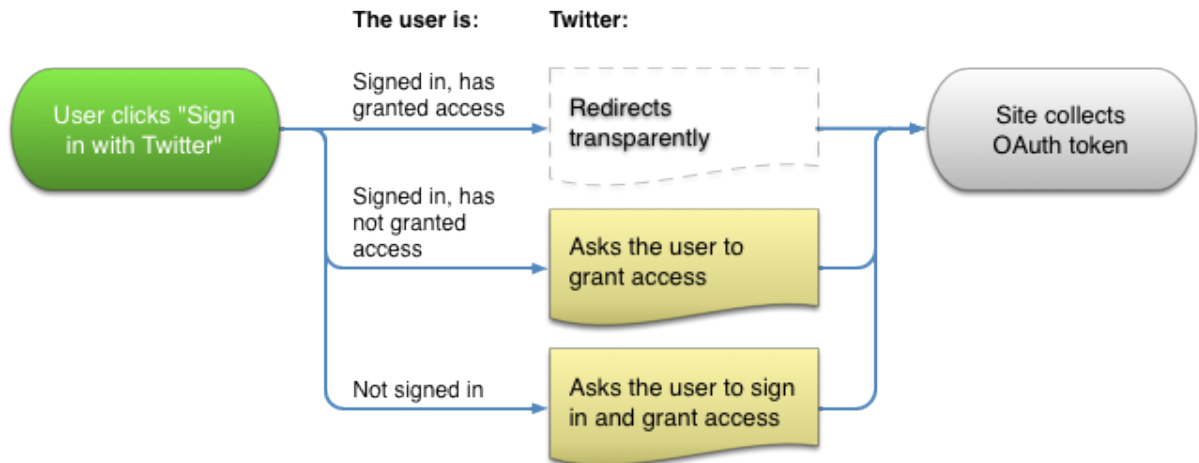


Figura 13 Flujo de solicitud de acceso de Twitter⁶⁶

- **Proceso de autenticación, Google+**

La red social **GOOGLE+** también se basa en el protocolo **OAuth 2.0**, y el proceso de autenticación es similar, la aplicación envía la petición de autorización, el proveedor genera el dialogo OAuth con la solicitud de permisos, a su vez, el usuario envía la respuesta, si el usuario acepta los permisos, se genera el **access token** necesario para acceder a los datos del usuario. La figura 14 muestra el flujo de petición de un **access token** en Google+.

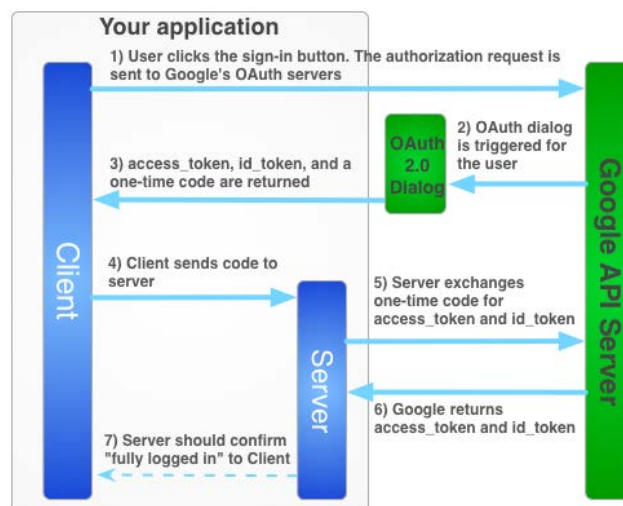


Figura 14 Flujo de solicitud de acceso de Google+⁶⁷

⁶⁶ <https://dev.twitter.com/docs/auth/implementing-sign-twitter>

⁶⁷ <https://developers.google.com/+/web/signin/server-side-flow>

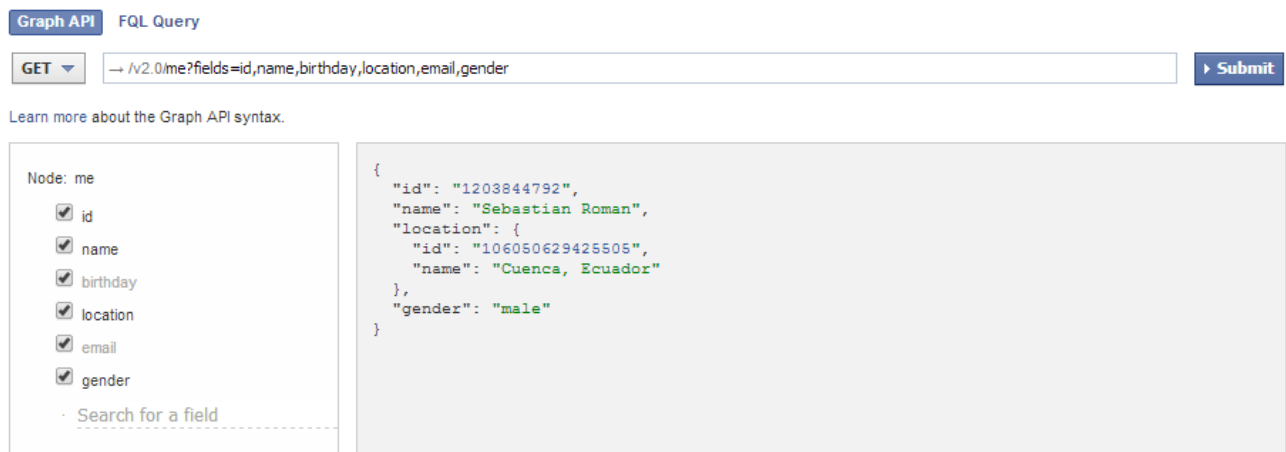
4.4.1.4. Formatos de respuesta

Cada *red social* tiene una *API* que maneja las solicitudes de datos del usuario, junto con el proceso de autenticación. Después de definir las *redes sociales* que serán la fuente de los datos implícitos, y luego de definir los procesos de autenticación, se detalla el **formato de los datos que contiene respuesta a la solicitud**. En la tabla **19** se detallan los diferentes *formatos* que soporta la respuesta de la solicitud.

RED SOCIAL	JSON	XML	Objeto	Comentarios
Facebook	No	No	Si	Facebook devuelve un objeto llamado GraphObject
Twitter	Si	No	No	Twitter devuelve una estructura de datos JSON
Google+	Si	No	No	Google devuelve una estructura de datos JSON

Tabla 19 Formato de respuesta de solicitud de las Redes sociales.

A continuación, las figuras **15**, **16** y **17** muestran pequeños ejemplos del formato de respuesta obtenido por Facebook, Twitter y Google+ respectivamente.



The screenshot shows the Facebook Graph API interface. At the top, there are tabs for 'Graph API' and 'FQL Query'. Below them is a text input field containing the URL 'v2.0/me?fields=id,name,birthday,location,email,gender' and a 'Submit' button. Below the input field, there is a link 'Learn more about the Graph API syntax.' On the left side, there is a section titled 'Node: me' with a list of fields: 'id', 'name', 'birthday', 'location', 'email', and 'gender', each with a checked checkbox. Below this list is a search bar labeled 'Search for a field'. On the right side, there is a large text area displaying the JSON response: { "id": "1203844792", "name": "Sebastian Roman", "location": { "id": "106050629425505", "name": "Cuenca, Ecuador" }, "gender": "male" }

Figura 15 Formato de respuesta de solicitud de Facebook



UNIVERSIDAD DE CUENCA

Example Request

GET https://api.twitter.com/1.1/account/verify_credentials.json

```
1. {
2.   "contributors_enabled": true,
3.   "created_at": "Sat May 09 17:58:22 +0000 2009",
4.   "default_profile": false,
5.   "default_profile_image": false,
6.   "description": "I taught your phone that thing you like. The Mobile Partner Engineer
   @Twitter. ",
7.   "favourites_count": 588,
8.   "follow_request_sent": null,
9.   "followers_count": 10625,
10.  "following": null,
11.  "friends_count": 1181,
12.  "geo_enabled": true,
13.  "id": 38895958,
14.  "id_str": "38895958",
15.  "is_translator": false,
16.  "lang": "en",
17.  "listed_count": 190,
18.  "location": "San Francisco",
19.  "name": "Sean Cook",
20.  "notifications": null,
21.  "profile_background_color": "1A1B1F",
22.  "profile_background_image_url":
   "http://a0.twimg.com/profile_background_images/495742332/purty_wood.png",
23.  "profile_background_image_url_https":
   "https://si0.twimg.com/profile_background_images/495742332/purty_wood.png",
24.  "profile_background_tile": true,
25.  "profile_image_url":
   "http://a0.twimg.com/profile_images/1751506047/dead_sexy_normal.JPG",
26.  "profile_image_url_https":
   "https://si0.twimg.com/profile_images/1751506047/dead_sexy_normal.JPG",
```

Figura 16 Formato de respuesta de solicitud de Twitter

Response

```
200 OK

- Show headers -

{
  -"emails": [
    -{
      "value": "abuelcore@gmail.com",
      "type": "account"
    }
  ],
  -"name": {
    "familyName": "Roman",
    "givenName": "Sebastian"
  },
  -"ageRange": {
    "min": 21
  }
}
```

Figura 17 Formato de respuesta de solicitud de Google+

4.4.2. MÓDULO DE REDES SOCIALES MRS 1.0

Después de seleccionar las redes sociales para el alcance del proyecto en la tabla 1 del capítulo 2, y también de observar los métodos de acceso a los datos del usuario en el punto 4.4.1.3, es necesario crear un **módulo unificado** para el proyecto. Según la estructura definida en el prototipo, el módulo a integrar los métodos de acceso a las redes sociales seleccionadas debe ser **independiente** de la aplicación web del proyecto. Además, debe **almacenar** los datos de acceso para usos futuros y debe permitir el **acceso** remotamente desde cualquier navegador web o computador.

Inicialmente se debería crear un módulo que reúna todas estas características, pero gracias a la librería **HYBRIDAUTH**⁶⁸ solo es necesario reutilizar y configurar esta librería debido a que cumple con las características necesarias para el proceso antes mencionado.

4.4.2.1. Librería HybridAuth

HybridAuth es una librería Open Source que se distribuye bajo doble licencia MIT y GPL⁶⁹, el objetivo principal de la librería es de actuar como una capa abstracta entre la **api** de la red social y la aplicación; está desarrollada en *PHP*, lo cual encaja perfectamente a la arquitectura del proyecto.

La librería permite desarrollar fácilmente aplicaciones o páginas web sociales, que implementen autenticación con las diferentes redes sociales, permitiendo el intercambio de perfiles de usuario, lista de amigos/contactos, línea de actividades actualizaciones de estado y más. Para el alcance del proyecto, la librería será utilizada para recuperar los datos relevantes del usuario como son su perfil, a quien sigue, sus gustos y sus actividades.

4.4.2.2. Configuración librería HybridAuth

Después de descargar la librería desde su página oficial⁷⁰, el primer paso para configurar la librería, es crear las aplicaciones correspondientes para cada red social, en sus respectivas páginas, con esto obtendremos dos valores:

APP ID que es el id de la aplicación creada en la red social, sirve como identificación de acceso.

APP SECRET que es una clave de acceso, sirve como nivel de seguridad para la autenticación.

⁶⁸ <http://hybridauth.sourceforge.net/>

⁶⁹ <http://hybridauth.sourceforge.net/licenses.html>

⁷⁰ <http://hybridauth.sourceforge.net/download.html>



UNIVERSIDAD DE CUENCA

Las figuras 18, 19, 20 muestran ejemplos de los respectivos paneles de las redes sociales donde se observan los parámetros **app id y app secret** mencionados.

Basic	Advanced	Migrations
<div>App ID</div> <div>225542074300087</div>		
<div>App Secret</div> <div>.....</div> <div>Show</div>		
<div>Display Name</div> <div>TvDigitalPrueba2</div>		
<div>Namespace</div> <div>tvdigitalprueba</div>		
<div>App Domains</div> <div></div>		
<div>Contact Email</div> <div>abuelcore@hotmail.com</div>		

Figura 18 Panel de administración de Facebook

Client ID for web application

Client ID	156020777658-42sdmaglvdrjig6iv8obi6inrpqk4g45.apps.googleusercontent.com
Email address	156020777658-42sdmaglvdrjig6iv8obi6inrpqk4g45@developer.gserviceaccount.com
Client secret	HhoQnC8FtBq51S7_TWkwyljS
Redirect URIs	http://localhost/MRS/hybridauth/?hauth.done=Google
Javascript Origins	http://localhost:80

Figura 19 Panel de administración de Google+

Application settings

Keep the "API secret" a secret. This key should never be human-readable in your application.

API key	7PAQI0dDwBygMolYRK4cFg
API secret	6GQ1O3Y5OISKJ8AIVfSl0s0YXfISljkLCpQ6zQg
Access level	Read-only (modify app permissions)
Owner	SebastianRomanJ
Owner ID	373621416

Figura 20 Panel de administración de Twitter

Una vez creado las aplicaciones, y generado los dos parámetros, se procede a configurar la librería, para eso debemos fijarnos en la estructura de archivos de la librería. Donde descargamos la librería, existe una carpeta: */hybridauth* en ella encontraremos el archivo *config.php*, como muestra la figura 21.

Hybrid	13/05/2014 10:29	Carpeta de archivos	
config.php	30/04/2014 9:35	PHP file	2 KB
error_auth.log	29/05/2014 16:18	Archivo LOG	475 KB
index.php	15/05/2014 1:44	PHP file	1 KB

Figura 21 Estructura de archivos librería HybridAuth

En el archivo configuramos el *array* de las redes sociales con los parámetros *APP ID* y *APP SECRET*, junto con el *endpoint* (que es la dirección *url* a donde redirige la api, después de la autenticación, en este caso el parámetro *base_url*) como se muestra en la figura 22.

```
array(
    "base_url" => "http://localhost/MRS/hybridauth/",

    "providers" => array (

        "Google" => array (
            "enabled" => true,
            "keys"    => array ( "id" => "156020777658-42sdmaglvdrjig6iv8obi6inrpqk4g45.apps.googleusercontent.com", "secret" => "GOCSPx-156020777658-42sdmaglvdrjig6iv8obi6inrpqk4g45.apps.googleusercontent.com" )
        ),

        "Facebook" => array (
            "enabled" => true,
            "keys"    => array ( "id" => "225542074300087", "secret" => "95d85d4398da899e7e2ef0da3326e469" )
        ),

        "Twitter" => array (
            "enabled" => true,
            "keys"    => array ( "key" => "IB75SpVvNGxqX4rK3wHHIw", "secret" => "aKVqHPIBf355umicVWzHmFeVucj" )
        ),

    ),
)
```

Figura 22 Configuración del array de redes sociales de la librería HybridAuth

Cuando se guarde el archivo, estará configurada la librería.

4.4.2.2.1. Estructura librería HybridAuth

La librería HybridAuth se compone de varias clases, cada una de ellas cumple una función específica en todo el módulo de redes sociales (Figura 23). A continuación veremos una breve descripción de cada una de ellas.⁷¹

- La clase **Hybrid_Auth** es la clase principal, provee una forma sencilla de autenticar a los usuarios a través de los protocolos OAuth y OpenID vistos en el punto 4.4.1.2. Esta es la clase que se debe instanciar para usar la librería.
- **Hybrid_Provider_Adapter** es la clase que sirve como un “puente” entre la librería y la API del proveedor (red social) determinado, se utiliza para conectar a los usuarios a una red social determinada.
- **Hybrid_Model** es la clase abstracta del modelo del proveedor específico (Facebook, Twitter o Google+), implementa los métodos de cada clase de las redes sociales: **Hybrid_Providers_Google**, **Hybrid_Providers_Twitter**, y **Hybrid_Providers_Facebook**.
- **Hybrid_User** es una clase que se utiliza para acceder a los datos del perfil de usuario: **Hybrid_User_Profile** desde la clase modelo. La clase que define el perfil del usuario contiene atributos que son genéricos para cualquier proveedor o red social.
- **Hybrid_Likes**, **Hybrid_User_Contact** y **Hybrid_User_Activity** son clases que se utilizan para acceder a los diferentes datos del usuario como sus intereses, sus contactos y sus actividades. Cada clase define atributos fijos que son genéricos para cualquier proveedor.
- **Hybrid_Error**, **Hybrid_Endpoint** y **Hybrid_Storage** son clases funcionales de la librería que ayudan al manejo de errores, el procesamiento de la dirección endpoint⁷², y manejo de sesiones respectivamente.
- **Hybrid_Logger**, es la clase que almacena en un archivo de texto plano los eventos que pasan paso a paso en la librería.

⁷¹ <http://hybridauth.sourceforge.net/apidoc.html>

⁷² El procesamiento del endpoint se refiere a formar la dirección url que permite el inicio de sesión del proveedor.

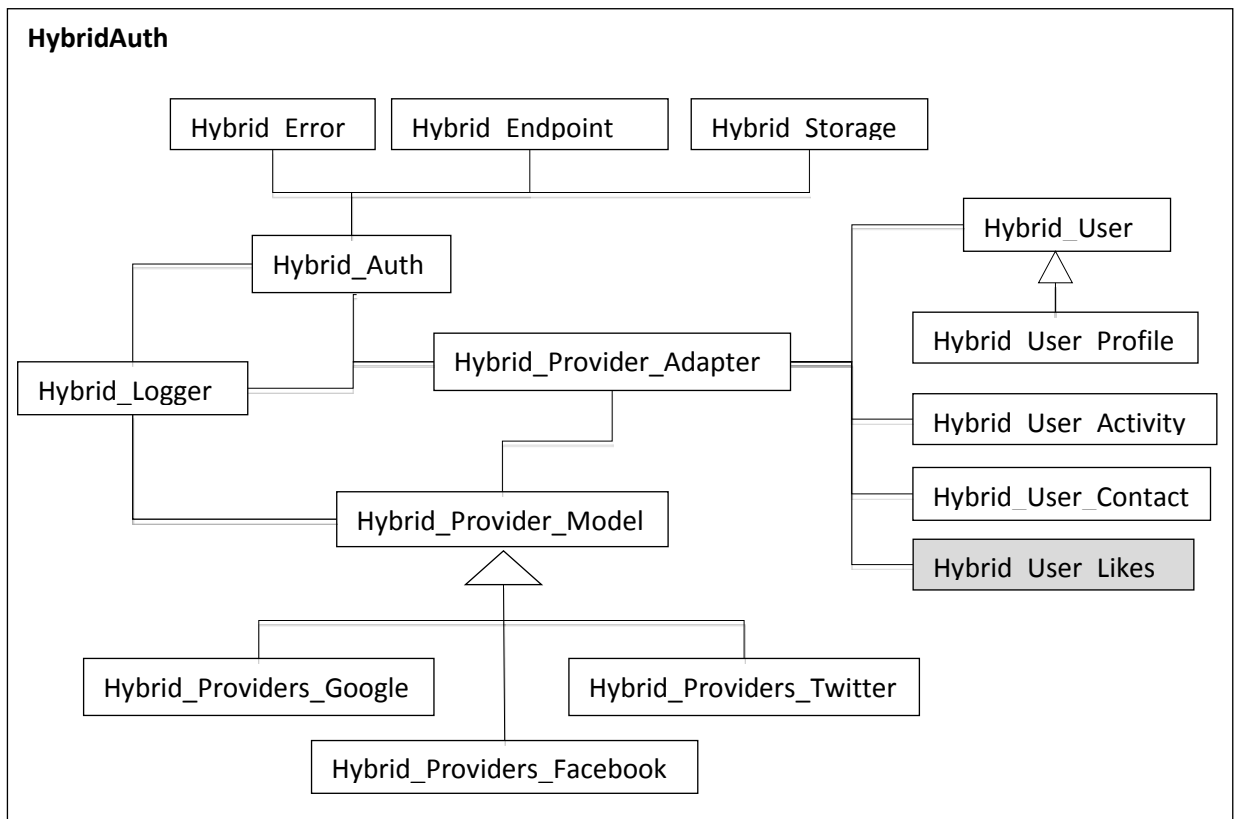


Figura 23 Estructura de clases librería HybridAuth⁷³

4.4.2.2.2. Métodos librería HybridAuth

La librería HybridAuth contiene diversos métodos en todas las clases vistas en el punto 4.4.2.2.2, así como las clases, los diferentes métodos tienen funcionalidades específicas para la librería. En esta sección se describirán los métodos más importantes.

- El método ***authenticate ()*** de la clase *hybrid_auth*, trata de autenticar al usuario con el proveedor dado, si el usuario ya está conectado con dicha red social, retorna una instancia de la clase *hybrid_adapter*, caso contrario trata de autenticar y autorizar al usuario con el proveedor.
- El método ***getAdapter ()*** de la clase *hybrid_auth*, retorna la instancia del usuario autenticado con el proveedor.
- El método ***redirect ()*** de la clase *hybrid_auth*, es una función útil que redirige a la url dada, tiene dos formas de hacerlo: por *PHP* o por *JavaScript*.
- El método ***getUserProfile ()*** de la clase *Hybrid_Provider_Adapter*, es una función que retorna el perfil del usuario autenticado en la librería.

⁷³ <http://hybridauth.sourceforge.net/apidoc.html>

- El método ***getUserContacts ()*** de la clase *Hybrid_Provider_Adapter*, es una función que retorna la lista de contactos/seguídos del usuario. No todos los proveedores soportan este método.
- El método ***getUserActivity ()*** de la clase *Hybrid_Provider_Adapter*, es una función que retorna la línea de actividad del usuario autenticado en la librería. No todos los proveedores soportan este método.
- El método ***getUserLikes ()*** de la clase *Hybrid_Provider_Adapter*, es una función que retorna los “likes” del usuario autenticado en la librería. No todos los proveedores soportan este método.
- El método ***logout ()*** de la clase *Hybrid_Provider_Adapter*, es una función que permite borrar todos los datos de la sesión **SOLO** del usuario actual.

4.4.2.2.3. Manejo de Sesiones de la librería HybridAuth

Para finalizar el análisis de la librería HybridAuth, es importante describir el manejo de **sesiones** que utiliza ⁷⁴, pues de esto depende mucho la autenticación.

Cuando un usuario se autentifica dando un proveedor, HybridAuth crea una **sesión** para el usuario, por defecto la sesión junto con sus datos se mantendrá siempre y cuando las actuales sesiones de *PHP* se mantengan activas o hasta que se llame al método *Hybrid_Provider_Adapter :: logout()*.

El usuario que se conecta a través de un proveedor determinado, utilizando HybridAuth, debe considerar 3 sesiones diferentes:

1. La primera sesión está en el lado del proveedor (Facebook, Twitter o Google+)
2. La segunda sesión es la que se crea en HybridAuth
3. Y la última sesión la que se utilice dentro de la aplicación.

Si la primera sesión se interrumpe por alguna razón (si las sesiones *PHP* expiran, o si el usuario revoca el permiso otorgado para la aplicación, mientras que él todavía tiene una sesión con HybridAuth, etc.), entonces HybridAuth ya no será capaz de comunicarse con la *API* del proveedor; y, en este caso se debe utilizar el método *Hybrid_Provider_Adapter :: logout ()* para que la librería elimine los datos de sesión del usuario, y se trate de autenticar al usuario de nuevo.

Este proceso llegaría a ser incómodo para el usuario que se registra así como para la aplicación que accede a la librería, es por eso que se necesita

⁷⁴ http://hybridauth.sourceforge.net/userguide/HybridAuth_Sessions.html

almacenar los datos de sesión del usuario para impedir que se autentifique cada vez.

En el punto 4.4.2.3.3 del presente capítulo se detallara los problemas de sesión que ocurren cuando se accede remotamente a la librería, y su solución.

4.4.2.3. API RestFull

Después de seleccionar la librería HybridAuth como base para el *módulo de redes sociales MRS*, **y considerando los requisitos planteados** en la sección 4.4.2, una de las características que debía tener el módulo era también permitir el **acceso** remoto, independiente de cualquier plataforma⁷⁵.

Existen muchas alternativas para acceder a aplicaciones desde lugares remotos: servicios SOAP, servicios REST, sockets, RPC, etc. En el proyecto se decidió utilizar como interfaz de acceso remoto una **API RestFull** para el módulo de redes sociales **MRS**, debido a las ventajas frente a los otros métodos, descritas en la tabla 20.

Ventajas	Comentario
Ofrece posibilidades de programar con menos código	Al ser orientado a recursos y stateless se necesita menos código para procesar solicitudes
Arquitectura escalable	Permite soportar nuevo contenido en el módulo sin impacto en la funcionalidad
Tiempo de respuestas más rápidas	Orientado a recursos no a procesos, lo que provoca respuestas más rápidas,
Consumo de pocos recursos	REST es particularmente útil en dispositivos con escasos recursos como PDAs o teléfonos móviles, donde la sobrecarga de las cabeceras y capas adicionales de otros métodos deben ser restringidas.
Fácil acceso	Al ser orientado a recursos y stateless provee mayor facilidad de acceso
Buen soporte	Los servicios REST, están teniendo

⁷⁵ El módulo MRS está configurado para ambiente **web**, es por eso que las aplicaciones que desean consumir los datos, deben estar en un ambiente web.

	bastante acogida en el medio
--	------------------------------

Tabla 20 Ventajas de servicios REST

Después de seleccionar la capa de acceso al módulo MRS, debemos especificar los métodos REST de acceso al módulo, y sus características principales.

4.4.2.3.1. Métodos API RestFull

La librería HybridAuth está desarrollada en *PHP*, debido a esto se debe encontrar un framework⁷⁶ que permita dar soporte a una arquitectura *REST* en *php*, es por eso que para el alcance del proyecto se ha decidido utilizar la librería **SLIM**⁷⁷ debido a sus características principales: ruteo rápido, configuración simple, gran soporte y utiliza pocos recursos.

Los métodos de acceso creados en la **API RestFull** del módulo **MRS** vienen especificados en la tabla 21.

URL	Método	Parámetros	Descripción
/islogged	GET	Provider, id	Devuelve una respuesta '1' si el proveedor del usuario id está configurado, o nulo si no
/authenticate	GET	Provider, url, id	Configura el modulo con el Provider especificado para el usuario id.
/getProfile	GET	Provider, id	Obtiene el perfil del usuario específico, id y el proveedor provider
/getLikes	GET	Provider, id	Obtiene los "likes" del usuario específico, id y el proveedor provider
/getContacs	GET	Provider, id	Obtiene los contactos del usuario específico, id y el proveedor provider
/getActivities	GET	Provider, id	Obtiene las actividades del usuario específico, id y el proveedor provider
/logout	GET	Provider	Termina sesión del usuario con el provider

Tabla 21 Métodos de acceso REST al módulo MRS

4.4.2.3.2. Formato de respuesta MRS

El formato de datos que devuelve la **API RestFull** del módulo **MRS** es único para cualquier solicitud, es decir cualquier método **GET** para obtener datos del usuario ya sea su perfil, contactos, likes y actividades será el mismo formato de respuesta. Para el alcance del proyecto se ha decidido utilizar un formato de

⁷⁶ Framework es una estructura conceptual y tecnológica de soporte definido, normalmente con artefactos o módulos de software concretos, que puede servir de base para la organización y desarrollo de software

⁷⁷ <http://www.slimframework.com/>



UNIVERSIDAD DE CUENCA

respuesta **JSON**. Las figuras 24, 25, 26 y 27 muestran el formato de dato de los diferentes métodos *REST*.

Raw	JSON	Response
Copy to clipboard Save as file		
<pre>{ http://schema.org/name: "Sebastian Roman" http://schema.org/email: "abuelcore@hotmail.com" http://schema.org/age: 24 http://schema.org/location: "Cuenca, Ecuador" http://schema.org/gender: "male" http://schema.org/occupation: "Computer Science" }</pre>		

Figura 24 Formato de respuesta JSON del perfil de usuario en Facebook

Raw	JSON	Response
Copy to clipboard Save as file		
<pre>[193] -0: { http://schema.org/likes/id: "213090532184572" http://schema.org/likes/created_time: "2014-06-05T16:28:38+0000" http://schema.org/likes/name: "Borrachera Futbolera Ecuador" http://schema.org/likes/category: "Community" } -1: { http://schema.org/likes/id: "261896753997599" http://schema.org/likes/created_time: "2014-06-04T16:41:52+0000" http://schema.org/likes/name: "Si, a la cadena perpetua a los asesinos en el Ecuador" http://schema.org/likes/category: "Community" } -2: { http://schema.org/likes/id: "632836833452607" http://schema.org/likes/created_time: "2014-06-04T16:37:50+0000" http://schema.org/likes/name: "FEF" http://schema.org/likes/category: "Organization" }</pre>		

Figura 25 Formato de respuesta JSON de los “likes” de usuario en Facebook

Raw	JSON	Response
Copy to clipboard Save as file		
<pre>[15] -0: { http://schema.org/contacs/identifiier: 282768438 http://schema.org/contacs/displaName: "Ambiente Ecuador" http://schema.org/contacs/profileURL: "http://twitter.com/Ambiente_Ec" http://schema.org/contacs/photoURL: "http://pbs.twimg.com/profile_images/245 http://schema.org/contacs/description: "Cuenta Oficial de Twitter del Minist } -1: { http://schema.org/contacs/identifiier: 939367022 http://schema.org/contacs/displaName: "Jorge Glas Espinel" http://schema.org/contacs/profileURL: "http://twitter.com/JorgeGlas" http://schema.org/contacs/photoURL: "http://pbs.twimg.com/profile_images/375 http://schema.org/contacs/description: "Vicepresidente de la República del E }</pre>		

Figura 26 Formato de respuesta JSON de los contactos de usuario en Twitter (a quien sigue)

Raw	JSON	Response
Copy to clipboard Save as file		
<pre>[6] -0: { http://schema.org/contacs/identifiier: "z13aw35bhxu3t52jg04cihxhqgm4wtt4bjw" http://schema.org/contacs/date: "2014-05-14T22:52:16.481Z" http://schema.org/contacs/description: "The Mamas & The Papas: California Dreamin'" http://schema.org/contacs/user: "Sebastian Roman" } -1: { http://schema.org/contacs/identifiier: "z125v1k45kuqi3eon04cihxhqgm4wtt4bjw" http://schema.org/contacs/date: "2014-03-28T17:57:35.776Z" http://schema.org/contacs/description: "¿Por qué los niños no quieren ser Científico: http://schema.org/contacs/user: "Sebastian Roman" } -2: { http://schema.org/contacs/identifiier: "z13gulzxqpb2ynikl23njfhpobaidlvz" http://schema.org/contacs/date: "2014-03-20T15:51:16.910Z" http://schema.org/contacs/description: "Doc Watson & Clarence Ashley - House of the I http://schema.org/contacs/user: "Sebastian Roman" } }</pre>		

Figura 27 Formato de respuesta JSON de las actividades del usuario en Google+

4.4.2.3.3. Problemas encontrados

Al momento de implementar el módulo de redes sociales **MRS**, se detectaron algunos problemas al momento de acceder remotamente por medio de la **API RestFull**, dichos problemas se detallan a continuación.

- **REST servicio stateless.** Al decir que una api RestFull es stateless, estamos diciendo que no maneja sesiones y estados de variables, pues bien, esto representó un problema ya que en el punto 4.4.2.1.4 del presente capítulo se aclaró que la librería HybridAuth utilizaba sesiones de **PHP** para la autenticación, entonces cuando se accede al módulo

desde un acceso remoto, las variables de sesión se reinician provocando un error al momento de procesar el endpoint.

Solución: la solución más viable para el problema, fue almacenar en una pequeña base de datos las sesiones del usuario, como se lo puede observar en la figura 28.

+ Opciones

key	value	type	id
php_session_id	s:26:"b9m9fa9kc7fvk12la9sqnr0v5";	HA::CONFIG	Pundush
version	s:5:"2.1.2";	HA::CONFIG	Pundush
php_session_id	s:26:"4nid8fn0o4164n8kit7oe8hpc6";	HA::CONFIG	Pundush
hauth_session.facebook.hauth_return_to	s:139:"http://localhost/MRS/REST/api/authenticate/...	HA::STORE	Pundush
hauth_session.facebook.hauth_endpoint	s:52:"http://localhost/MRS/hybridauth/?hauth.done=...	HA::STORE	Pundush
hauth_session.facebook.id_provider_params	a:5:{s:15:"hauth_return_to";s:139:"http://localhos...	HA::STORE	Pundush

Figura 28 Estructura tabla sesiones HybridAuth

- **REST orientado a recursos, sin re-direccionamiento.** Otro problema fue la autenticación del usuario. Un servicio REST es orientado a recursos, se realiza una petición de un recurso, y la *API* devuelve una respuesta; no existen pasos intermedios.

Uno de los principales problemas es que el protocolo de autenticación OAuth descrito en el punto 4.4.1.2. del presente capítulo, tiene varios pasos intermedios cuando se realiza la autenticación por primera vez, así que el usuario cuando se intente autenticar por primera vez en el módulo de redes sociales, la respuesta del módulo será el **endpoint** para el re-direccionamiento, y no el recurso solicitado, este problema está ilustrado en la figura 29.

Raw	Parsed	Response
Open output in new window Copy to clipboard Save as file Open in JSON tab		
<pre><html><head><script type="text/javascript">function redirect() { window.top.location.href="http://locall</pre>		

Figura 29 Respuesta del Módulo a una solicitud REST por primera vez

Solución: la solución a este problema fue generar los siguientes métodos:

/islogged

/authenticate

El primer método ***islogged***, permite saber si el usuario ya ha sido autenticado con el proveedor específico, si ya ha sido autenticado, no existe problema y se puede obtener los datos, pero si es la primera vez que accede al módulo con el proveedor específico, se llama al método ***authenticate*** accediendo directamente a la página del módulo, para permitir al protocolo OAuth realizar el proceso de autorización, y poder configurar el modulo para el usuario y proveedor específico.

4.4.3. Eventos Televisión Digital

Otra fuente importante para la recolección de datos implícitos del usuario, es la interacción con la televisión digital, los datos como programa, fecha, canal, y duración son importantes para alimentar el modelo de perfil de usuario para la televisión digital.

Para el acceso a estos datos, se utiliza un medio de almacenamiento intermedio entre la aplicación y la televisión digital, la figura 30 muestra el acceso a dichos datos, primero la televisión genera un conjunto de interacciones que tiene el usuario particular A, dicho usuario debe estar registrado tanto en la aplicación como en la televisión digital, después de almacenar los datos de la televisión digital, la aplicación recolecta la información del usuario particular A y lo almacena en la ontología.

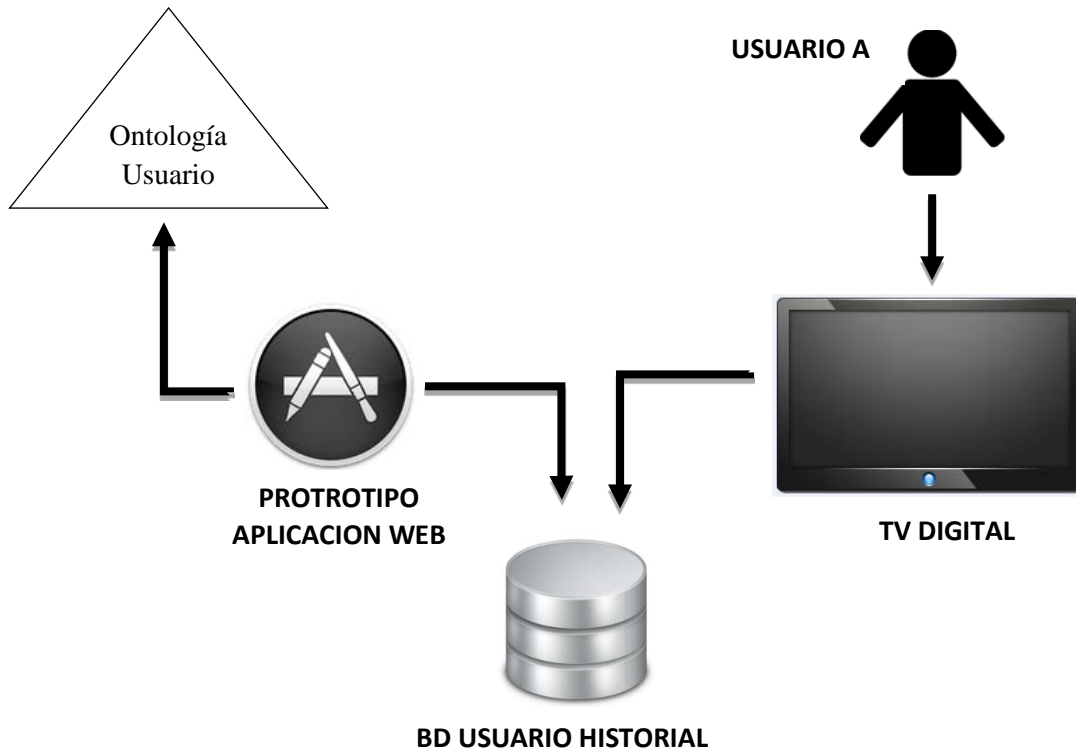


Figura 30 Acceso a los datos de la televisión digital

Para el alcance del proyecto de tesis se tomó en cuenta los eventos de la televisión digital que son recolectados desde una base de datos de eventos que fueron capturados en pruebas externas a esta tesis, pero que son parte del proyecto al que la tesis pertenece. Se realiza este procedimiento debido a que todavía no existe un mecanismo concreto que permita la comunicación entre la plataforma de televisión digital y aplicaciones web o aplicaciones externas a la televisión digital.



UNIVERSIDAD DE CUENCA

PROTOTIPO DE PERFIL DE USUARIO

A continuación se presentará el proceso de construcción del prototipo que gestiona el perfil de usuario con tecnologías semánticas.

CAPÍTULO 5

PROTOTIPO PERFIL DE USUARIO

En el presente capítulo se presentará detalles de la construcción del prototipo que gestiona los perfiles de usuario con tecnologías semánticas. El objetivo de este capítulo es primero presentar el proceso de diseño de la arquitectura del prototipo, después mostrar la configuración del repositorio semántico a ser usado, siguiendo con el proceso de análisis del software implementado. Una vez definida la arquitectura, se describe el proceso de implementación para luego presentar las pruebas realizadas y sus resultados.

5.1. ANÁLISIS

El prototipo propuesto busca diseñar y poblar un perfil de usuario con tecnologías semánticas utilizando alimentación de datos por parte del usuario y alimentándolo además con datos de redes sociales e interacciones con la televisión digital seleccionadas por sí mismo. El objetivo final del sistema es la alimentación de datos a un perfil de usuario basado en ontologías.

5.1.1. Funciones del prototipo

El prototipo debe cumplir las siguientes funciones básicas:

- Realizar la autenticación de los usuarios.
- Registrar usuarios desde un formulario en la web.
- Registrar a usuarios accediendo a sus redes sociales.
- Recolectar datos de un usuario desde un formulario de ingreso en la web y accediendo a sus redes sociales.
- Manejar preferencias de un usuario.
- Recolectar datos de la televisión digital desde la base de datos.
- Usar el motor de inferencia de la ontología para la toma de datos específicos

5.1.2. Escenarios base del prototipo

A continuación se detallarán los tres escenarios base que tendrá el sistema:

- En el primer escenario el usuario se registrará al sistema llenando el formulario de registro base en la web y posteriormente realizará el ingreso al sistema para revisar sus preferencias. Posteriormente, el usuario hará búsquedas sobre los datos ingresados.
- En el segundo escenario el usuario realizará el registro en el prototipo mediante una red social, es decir dará los permisos a la aplicación para ofrecer datos de sí mismo a través de su red social o redes sociales seleccionadas. Luego el usuario realizará búsquedas sobre los datos ingresados.
- Como tercer escenario el usuario realizará su registro mediante el formulario de registro en la web y posteriormente dará permisos a la aplicación para acceder a su perfil de red social o redes sociales. En la aplicación podrá ingresar más datos de su perfil, o también podrá obtener los datos provenientes de la televisión digital. Como siguiente paso el usuario hará consultas sobre los datos ingresados.

5.1.3. Identificación de actores

Una vez especificados los escenarios, se puede observar que el sistema tiene un solo actor que es el usuario de la aplicación. El mismo actor será el encargado de manejar todas las funcionalidades del sistema. El usuario podrá ingresar sus datos, buscarlos, editarlos o borrarlos de acuerdo a su necesidad.

5.1.4. Casos de uso del sistema

Se identificaron los siguientes casos de uso del sistema, obtenidos a partir de los escenarios presentados:

5.1.4.1. Caso de uso CU-01 Autenticación de Usuario.

El siguiente caso de uso está especificado en la tabla **22** donde se representa los detalles principales de la interacción del usuario cuando se autentifica en el prototipo de la aplicación web, se especifica la secuencia, los actores y las excepciones que permiten identificar la interacción del usuario con el prototipo.

CU- 01	Autenticación de Usuario	
Versión	1.0	
Autores	Román Sebastián Sáenz Juan José	
Descripción	El caso de uso base a la autenticación de un usuario donde presentará la petición de los datos necesarios para autenticarse en el sistema.	
Precondición	Ninguna.	
Secuencia Normal	Paso	Acción
	1.1.1	El sistema presenta pantalla de autenticación para usuarios (usuario y contraseña). Además el sistema presenta la opción de registro.
	1.1.2	El usuario ingresa los datos necesarios para autenticarse.
	1.1.3	El usuario presiona en la opción de login.
	1.1.4	El sistema verifica los datos ingresados por el usuario y le permite el ingreso al sistema.
Excepciones	Paso	Acción
	1.2.1	El usuario no llena los campos de autenticación, el sistema presenta un error y regresa al paso 1.1.1.
	1.2.2	El usuario llena campos con datos errados, el sistema presenta un error y regresa al paso 1.1.1.
Disparador	Validaciones de datos	
Frecuencia esperada	Siempre	
Importancia	Alta	
Comentarios	Se verifica los datos de usuario en la base de datos del prototipo.	

Tabla 22 Especificación de caso de uso: Autenticación de Usuario

5.1.4.2. Caso de uso CU-02 Registro de Usuario mediante formulario

La tabla 23 representa los detalles principales de la interacción del usuario cuando se registra utilizando un formulario web, la secuencia, los actores y las excepciones permiten identificar la interacción del usuario con el prototipo.

CU- 02	Registro de usuario mediante formulario
Versión	1.0
Autores	Román Sebastián Sáenz Juan José



Descripción	El caso de uso base de registro de un usuario donde presentará la petición de los datos necesarios para completar un perfil mínimo.	
Precondición	Ninguna.	
Secuencia Normal	Paso	Acción
	1.1.1	El sistema presenta una pantalla de autenticación para usuarios (usuario y contraseña). Además el sistema presenta la opción de registro.
	1.1.2	El usuario ingresa a la opción de registrarse.
	1.1.3	El sistema presenta la opción de registro con redes sociales o formulario.
	1.1.4	El usuario utiliza el registro mediante formulario.
	1.1.5	El sistema presenta el formulario con los datos a ser llenados por parte del usuario.
	1.1.6	El usuario presiona en la opción de grabar los datos.
	1.1.7	El sistema almacena los datos ingresados por el usuario.
Excepciones	Paso	Acción
	1.2.1	El usuario no llena los campos obligatorios de registro, el sistema presenta un error y regresa al paso 1.1.5.
	1.2.2	El usuario llena campos con datos errados, el sistema presenta un error y regresa al paso 1.1.5.
Disparador	Validaciones de datos	
Frecuencia esperada	Siempre	
Importancia	Alta	
Comentarios	Se registra los datos de perfil de usuario en la ontología seleccionada.	

Tabla 23 Especificación de caso de uso: Registro de usuario mediante formulario Web

5.1.4.3. Caso de uso CU-03 Registro de Usuario mediante redes sociales

El siguiente caso de uso está especificado en la tabla **24** donde se representa los detalles principales de la interacción del usuario cuando se registra utilizando redes sociales, la secuencia, los actores y las excepciones permiten identificar la interacción del usuario con el prototipo.

CU- 03	Registro de Usuario mediante redes sociales
Versión	1.0
Autores	Román Sebastián Sáenz Juan José



Descripción	El caso de uso base de registro de un usuario donde tomará los datos desde la red social especificada.	
Precondición	Ninguna.	
Secuencia Normal	Paso	Acción
	1.1.1	El sistema presenta una pantalla de autenticación para usuarios (usuario y contraseña). Además el sistema presenta la opción de registro.
	1.1.2	El usuario ingresa a la opción de registrarse.
	1.1.3	El sistema presenta la opción de registro con redes sociales o formulario.
	1.1.4	El usuario utiliza el registro mediante redes sociales.
	1.1.5	El sistema presenta las redes sociales disponibles.
	1.1.6	El usuario selecciona la red social
	1.1.7	El usuario otorga permisos de la red social
	1.1.8	El sistema almacena los datos obtenidos de la red social.
Excepciones	Paso	Acción
	1.2.1	El usuario no acepta los permisos de la red social y el sistema regresa al paso 1.1.3.
Disparador	Validaciones de datos	
Frecuencia esperada	Siempre	
Importancia	Alta	
Comentarios	Se registra los datos de perfil de usuario en la ontología seleccionada.	

Tabla 24 Especificación de caso de uso: Registro de usuario mediante redes sociales

5.1.4.4. Caso de uso CU-04 Registro de Usuario mediante formulario y redes sociales

El mejor escenario para el prototipo está especificado en el caso de uso descrito en la tabla 25 donde el usuario decide registrarse utilizando las dos alternativas: formulario web y redes sociales. Los detalles de la secuencia, los actores y las excepciones permiten identificar la interacción del usuario con el prototipo.

CU- 04	Registro de Usuario formulario y permiso de redes sociales
Versión	1.0



UNIVERSIDAD DE CUENCA

Autores	Román Sebastián Sáenz Juan José	
Descripción	El caso de uso base de registro de un usuario donde tomará los datos de un formulario y lo complementará con una red social especificada.	
Precondición	Ninguna.	
Secuencia Normal	Paso	Acción
	1.1.1	El sistema presenta pantalla de autenticación para usuarios (usuario y contraseña). Además el sistema presenta la opción de registro.
	1.1.2	El usuario ingresa a la opción de registrarse.
	1.1.3	El sistema presenta la opción de registro con redes sociales o formulario.
	1.1.4	El usuario utiliza el registro mediante formulario y sigue el proceso desde CU-02 1.1.5 hasta CU-02 1.1.7.
	1.1.5	El sistema presenta la pantalla perfil.
	1.1.6	El usuario selecciona la red social en la pantalla de perfil para actualizar datos.
	1.1.7	El usuario elije si desea actualizar sus datos personales de la red social escogida.
	1.1.8	El sistema y sigue el proceso desde CU-03 1.1.6 hasta CU-03 1.1.8.
Excepciones	Paso	Acción
	1.2.1	El usuario no acepta los permisos de la red social y el sistema regresa al paso 1.1.5.
Disparador	Validaciones de datos	
Frecuencia esperada	Siempre	
Importancia	Alta	
Comentarios	Se registra los datos de perfil de usuario en la ontología seleccionada.	

Tabla 25 Especificación de caso de uso: Registro de usuario mediante formulario web y redes sociales

5.1.4.5. Caso de uso CU-05 Modificación de preferencias del usuario

La tabla **26** representa los detalles principales de la interacción del usuario cuando se añade, modifica o elimina sus preferencias utilizando el prototipo.

CU- 05	Modificación de preferencias del usuario
Versión	1.0



Autores	Román Sebastián Sáenz Juan José	
Descripción	El caso de uso que detallará el proceso de modificación de datos de un perfil almacenado.	
Precondición	Ninguna.	
Secuencia Normal	Paso	Acción
	1.1.1	El sistema presenta una pantalla de autenticación para usuarios (usuario y contraseña). Además el sistema presenta la opción de registro.
	1.1.2	El usuario realiza la identificación al sistema
	1.1.3	El sistema presenta la pantalla donde están las preferencias del usuario almacenadas.
	1.1.4	El usuario navega entre las preferencias y selecciona las que quiera eliminar o editar.
	1.1.5	El usuario añade/edita preferencias o las elimina.
	1.1.6	El sistema advierte con mensajes sobre los cambios realizados.
	1.1.7	El usuario acepta los cambios.
Excepciones	Paso	Acción
	1.2.1	EL usuario no tiene datos de preferencia almacenados, el sistema continua en el paso 1.1.3
Frecuencia esperada	Siempre	
Importancia	Alta	
Comentarios	Se registra los datos de perfil de usuario en la ontología seleccionada.	

Tabla 26 Especificación de caso de: **Modificación de preferencias del usuario**

5.2. DISEÑO

En todo proyecto de software es muy importante la fase de diseño, facilita la comprensión del prototipo y ayuda a construir e implementar los diferentes casos de uso que se definen en la fase de análisis.

El diseño del prototipo especifica la arquitectura interna del sistema, de que módulos se componen y la comunicación entre estos. Al ser un prototipo, el

objetivo del proyecto no busca crear sistemas complejos, la funcionalidad total de la aplicación está en base a interfaces básicas.

5.2.1. Arquitectura interna

La estructura del prototipo expuesta en el presente capítulo, contiene los siguientes módulos (Figura 31): MPU (Aplicación web) que contiene los *formularios web*, MRS (módulo de redes sociales) modulo especificado en el capítulo 4, MIT (módulo de ingreso de tripletas detallado en el capítulo 3 sección 3.3.2 "*población inicial de la ontología*" en [54]) y el repositorio semántico que contiene el modelo de usuario de televisión digital descrito en el presente capítulo, punto 5.3.

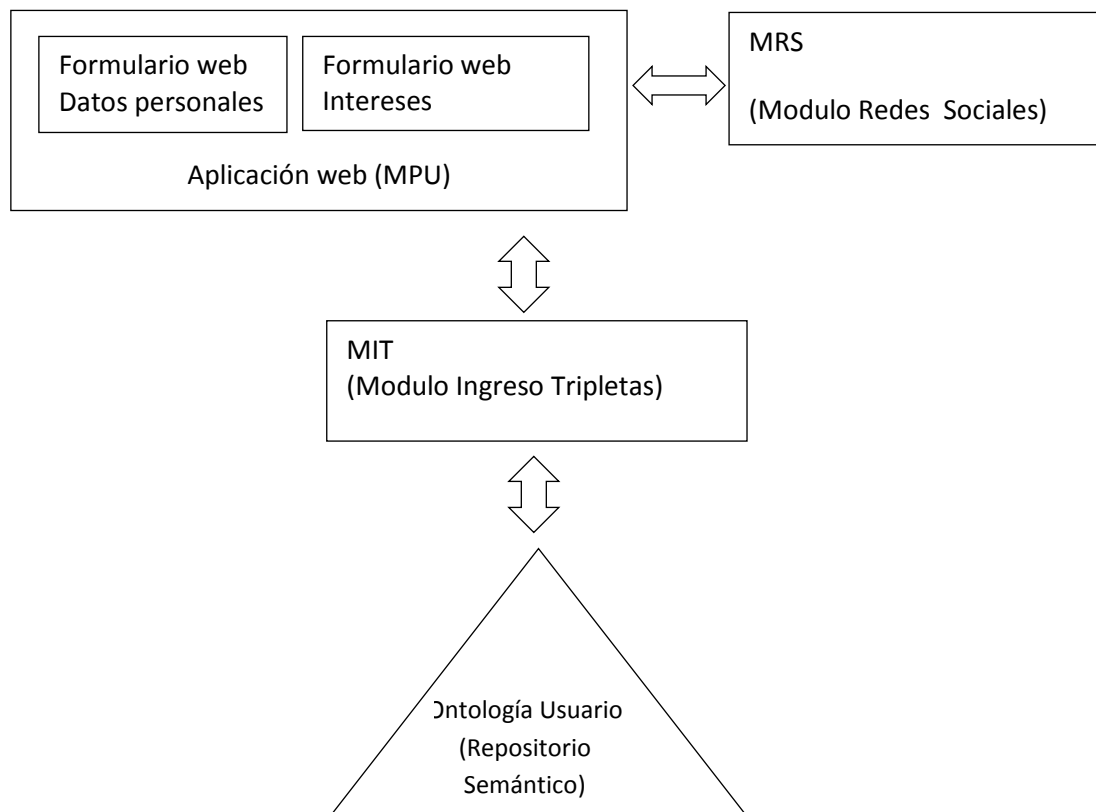


Figura 31 Diseño del Prototipo de modelamiento de usuario para la televisión digital

Los formularios web y el módulo *MRS* estarán desarrollados bajo la misma plataforma, a pesar de ser totalmente independientes, debido a la facilidad de

desarrollo. El repositorio que alojará la *Ontología Usuario* se describe a continuación.

5.3. REPOSITORIO

Para almacenar datos semánticos en una ontología, es necesario tener un almacén de datos que soporte la tecnología ontológica. Para ello se utiliza los repositorios semánticos como un medio de almacenamiento de los datos semánticos y una posterior consulta de los datos almacenados. En el desarrollo de la aplicación prototipo que se plantea en este proyecto es necesario utilizar un repositorio semántico que sirva como almacén de datos. Por este motivo a continuación se describirá un repositorio semántico y posteriormente se presentarán ejemplos de repositorios que se encuentran en el mercado.

Al referir a un repositorio semántico, se pretende definir a una aplicación de almacenamiento de recursos semánticos y sus metadatos, manejando una interfaz de búsqueda similar a un servidor de base de datos pero con la principal característica de almacén semántico con datos y metadatos. De esta manera, al almacenar datos y metadatos permite la búsqueda más rápida y optima de recurso e información similar en los objetos almacenados dentro del repositorio semántico. Además, el repositorio puede ser consultado mediante búsquedas por software o mediante interfaces para el usuario que genere ayudas y facilidades de consulta para la persona que lo necesite.

La variedad de repositorios semánticos y sus características son extensas tomando en consideración los diversos usos que tendrán y la alta demanda de características que generan las aplicaciones que posteriormente consumirán estos repositorios. A continuación se describirá de forma breve algunos de los repositorios semánticos disponibles en el mercado, y al final se indicará el repositorio seleccionado para el prototipo.

5.3.1. SESAME

Sesame es un repositorio semántico desarrollado por ADUANA que permite el almacenamiento de RDF con consultas en lenguajes RQL, RDQL, SPARQL y SeRQL, su principal objetivo es simplificar a un repositorio semántico orientando su arquitectura a RDF.

5.3.2. APACHE JENA

Creado por Apache Software Foundation, es un repositorio semántico que utiliza como principal característica el almacenamiento persistente y en memoria además de tener un motor para inferencia muy estable con características muy completas. La estabilidad del repositorio ayuda a la generación de aplicaciones con características de complejidad elevada sin mayores problemas

5.3.3. VIRTUOSO

El repositorio semántico VIRTUOSO⁷⁸, creado por OpenLink Software, tiene como principal característica la integración de SQL, RDF, XML y además servicios web. La integración de estos servicios permite la múltiple utilización del repositorio que además genera una arquitectura compuesta con la posibilidad de almacén de diversos datos.

Para describir de mejor manera la capacidad del repositorio VIRTUOSO, la figura 32 presenta la arquitectura del repositorio, donde se muestra la integración y diversidad de tecnologías dentro del mismo repositorio.

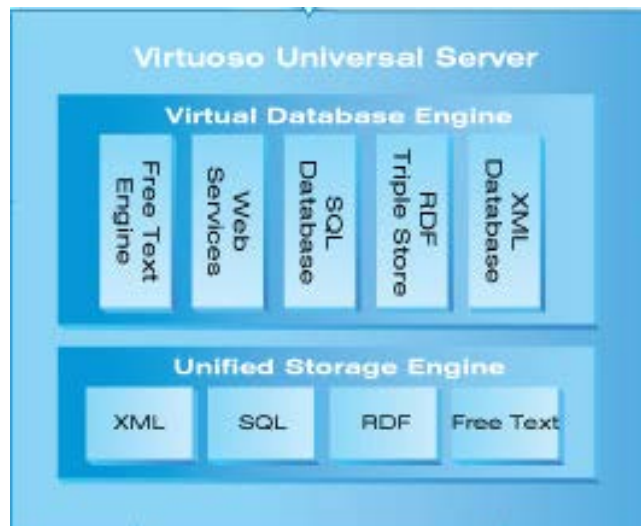


Figura 32 Arquitectura de VIRTUOSO. FUENTE: [53]

⁷⁸ <http://virtuoso.openlinksw.com/>
Román Jarrín Sebastián Rodrigo
Sáenz Peñafiel Juan José

Para el alcance del proyecto, se ha decidido utilizar el repositorio **VIRTUOSO** para dar soporte a la ontología. Las ventajas del repositorio se detallan en la tabla 27⁷⁹.

Ventajas	Comentario
Agilidad y Rapidez	Virtuoso permite mezclar los mejores servicios de gestores de base de datos, con una infraestructura ligera y fácil de acceder.
Reducción de costos	Debido a su licencia GPL reduce costos asociados a almacenar datos propietarios de aplicaciones.
Mayor agilidad	Capacidad de mezclar y combinar los diferentes datos mediante estándares actuales como HTTP, SPARQL, y SQL.
Permite acceso por servicios web	Esta característica es importante para el manejo del repositorio remotamente
Versatilidad	Virtuoso tiene la capacidad para encajar de forma natural y segura hacia datos a través de fuentes internas y externas

Tabla 27 Ventajas del repositorio VIRTUOSO

A continuación se detalla las principales funciones del repositorio **VIRTUOSO** utilizadas en el prototipo, dichas funciones son: ejecución local del servidor **VIRTUOSO**, ejecución de entorno web en el servidor **VIRTUOSO**, carga de modelo ontológico al repositorio **VIRTUOSO** y por último el acceso remoto al repositorio **VIRTUOSO**.

5.3.3.1. Ejecución local del servidor VIRTUOSO

Después de descargar el repositorio *Virtuoso* de su página oficial ⁸⁰, procedemos a instalar el software. Para poder utilizar *virtuoso* se debe instalar con la licencia con la que se descarga el instalador, una vez terminado de instalar, se ejecuta el servicio “Openlink Virtuoso Server”.

⁷⁹ [http://www.openlinksw.com/dataspace/doc/kidehen@openlinksw.com/weblog/kidehen@openlinksw.com/s%20BLOG%20\[127\]/1609](http://www.openlinksw.com/dataspace/doc/kidehen@openlinksw.com/weblog/kidehen@openlinksw.com/s%20BLOG%20[127]/1609)

⁸⁰ <http://virtuoso.openlinksw.com/download/>

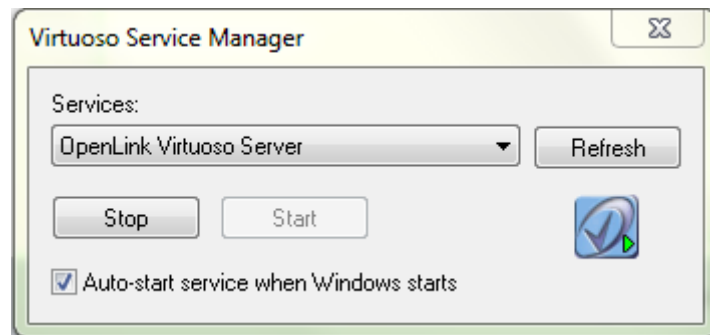


Figura 33 Interfaz de ejecución del repositorio VIRTUOSO

Si el servidor se ejecuta correctamente, para poder utilizar el entorno web del software virtuoso, se necesita abrir la siguiente dirección url en el navegador web: <http://localhost:8890>⁸¹, esta dirección está formada por la dirección **ip** del servidor (en este caso, para casos de prueba *localhost*) y el puerto 8890 por donde se escucha el servidor. La figura 34 muestra la interfaz de usuario del repositorio Virtuoso.



Figura 34 Entorno web del repositorio VIRTUOSO

Para poder utilizar los servicios del repositorio, se debe acceder a la pestaña **Conductor** en el menú lateral de la interfaz y autenticarse con el nombre de usuario: **dba** y contraseña: **dba (Usuario y contraseña por defecto)**.

⁸¹ El puerto 8890 es el puerto por defecto del servidor
Román Jarrín Sebastián Rodrigo
Sáenz Peñafiel Juan José

5.3.3.2. Entorno web servidor VIRTUOSO

Después de autenticarse en el servidor *Virtuoso*, la interfaz web del repositorio se habilita para el usuario por defecto **dba**. Para el alcance del proyecto, se utilizará el motor de gestión de tripletas del software, para eso, debemos acceder a la pestaña **Linked Data** de la interfaz web, cuando se abre el servicio, automáticamente se muestran todas las opciones para el manejo de tripletas, como se muestra la figura 35.

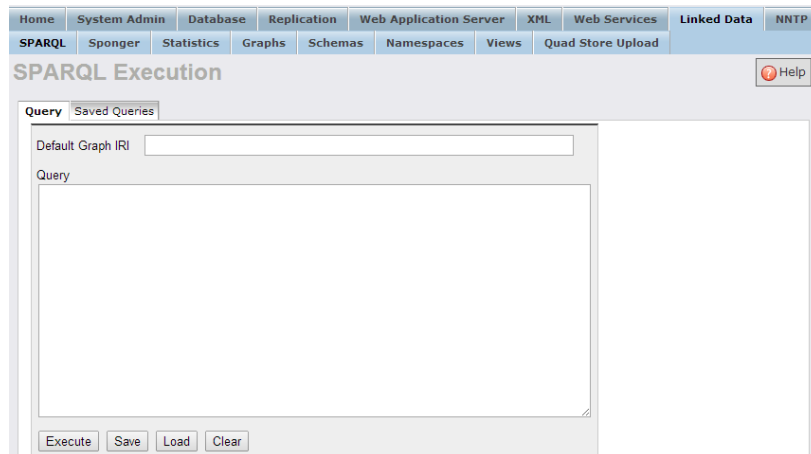


Figura 35 Interfaz Web “Linked Data” del repositorio VIRTUOSO

5.3.3.3. Carga de modelo ontológico al repositorio Virtuoso

Después de ejecutar el entorno web del repositorio, el cual nos permite gestionar tripletas de información, se procede a cargar el modelo ontológico del perfil de usuario para la televisión digital, especificado en el capítulo 3.

Para esto procedemos a dar clic en la pestaña **Quad Store Upload** del entorno web definido en el punto anterior. En ese momento aparecerá una ventana que nos permite cargar la ontología de perfil de usuario, seleccionando el archivo de texto plano (los archivos deben ser solo de tipo RDF, Turtle) especificando el nombre o IRI que identifica el grafo donde se cargara el modelo. Cuando se hayan especificado estos parámetros se procede a grabar el modelo, dando clic en **Upload**.



UNIVERSIDAD DE CUENCA

Home	System Admin	Database	Replication	Web Application Server	XML	Web Services	Linked Data
SPARQL	Sponger	Statistics	Graphs	Schemas	Namespaces	Views	Quad Store Upload

Quad Store Upload

☒ File* 20100809.rdf

☐ Resource URL*

☐ Create graph explicitly

Named Graph IRI*

Figura 36 Interfaz para cargar una ontología en el repositorio VIRTUOSO

Podemos confirmar que se ha cargado bien el modelo, accediendo a los grafos en la pestaña **Graphs** como muestra la figura 37.

Home	System Admin	Database	Replication	Web Application Server	XML	Web Services	Linked Data	NNTP
SPARQL	Sponger	Statistics	Graphs	Schemas	Namespaces	Views	Quad Store Upload	

Graphs Management

Graphs Users Security Roles Security Security Audit

Graph	Action
http://www.openlinksw.com/schemas/virtrdf#	Rename Delete
http://localhost:8890/sparql	Rename Delete
http://localhost:8890/DAV/	Rename Delete
http://www.w3.org/2002/07/owl#	Rename Delete
http://ucuenca.edu.ec/foaf/	Rename Delete
http://ucuenca.edu.ec/bbc/	Rename Delete
http://ucuenca.edu.ec/efoaf/	Rename Delete
http://ucuenca.edu.ec/perfil/	Rename Delete

Copyright © 1998-2014 OpenLink Software

Figura 37 Ontologías cargadas en el repositorio VIRTUOSO

5.3.4 Acceso remoto al repositorio VIRTUOSO

Una parte importante del proyecto, es el acceso remoto al repositorio, para poder ingresar, listar, modificar o eliminar tripletas del modelo de perfil de usuario ontológico, La figura 38 muestra en color gris el modulo que accede remotamente mediante un servicio REST al repositorio para el ingreso de tripletas, este servicio se encuentra detallado en el capítulo 3 sección 3.3.2 "población inicial de la ontología" en [54].

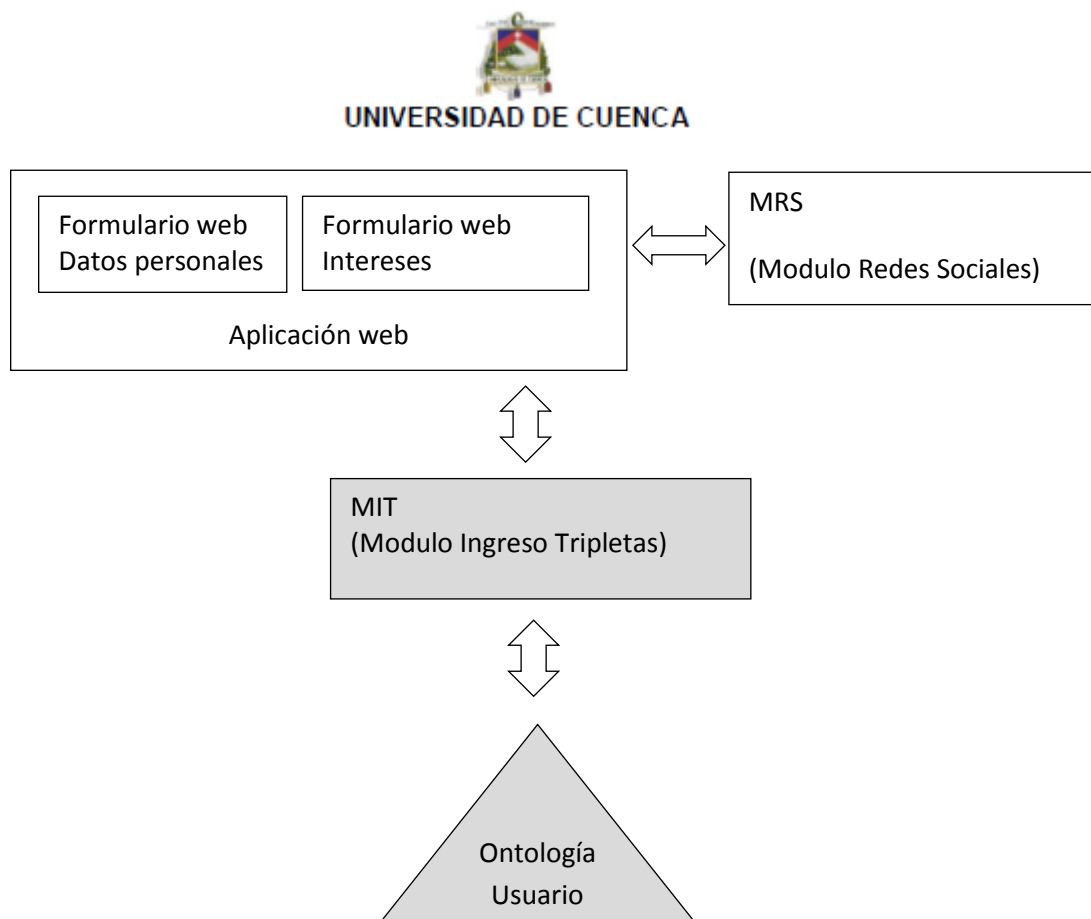


Figura 38 Acceso al Repositorio Virtuoso mediante el módulo MIT

Después de definir el repositorio y cargar el modelo de datos, a continuación se describe la fase de implementación del prototipo para el modelamiento del perfil de usuario utilizando tecnologías semánticas.

5.4. IMPLEMENTACIÓN

5.4.1. Plataforma de desarrollo web

Para la implementación del prototipo, se deben especificar características base, como el servidor donde se alojará la aplicación y el lenguaje de programación. La tabla **28**, detalla la lista de servidores de aplicación/web que existen en la actualidad.

Servidor	Tipo	Lenguajes que soporta	Licencia
Apache	Servidor Web	PHP/Perl/Python/Ruby/Asp	Libre - Licencia Apache 2.0
Apache Tomcat	Servidor Web/ Java Servlet	Servlets/JSPs	Libre - Licencia Apache 2.0
IIS	Servidor Web HTTP	ASP/ASP.NET/PHP/PERL	Software

			propietario - Microsoft Windows
Cherokee	Servidor web	PHP/Perl/Python/Ruby	Libre – GPL
Jboss	Servidor de aplicaciones	J2EE	Libre- LGPL
Glassfish	Servidor de Aplicaciones	J2EE	Libre – GPL

Tabla 28 Tabla comparativa, diferentes servidores web/aplicaciones⁸²⁸³⁸⁴⁸⁵⁸⁶⁸⁷

En la tabla **29** se comparan los diferentes servidores web y de aplicaciones que podrían dar soporte la *aplicación web* y por ende al *formulario web*. Para el alcance del proyecto, se ha decidido utilizar el servidor web **APACHE** debido a las ventajas descritas en esa tabla.

Ventajas	Comentario
<i>Experiencia de uso</i>	Existe un cierto nivel de experiencia con el uso del servidor Apache por parte de los desarrolladores, específicamente con el lenguaje de programación PHP, lo que llega a ser apropiado para la facilidad de realizar el proyecto
<i>Portabilidad</i>	El servidor Apache es una de los más portables ya que se es liviano y puedo instalarse en una amplia variedad de servidores y sistemas operativos
<i>Facilidad de uso</i>	El servidor web Apache es una de los más utilizados en la web, debido a su fácil uso y herramientas que dispone.
<i>Funcionalidades</i>	A pesar de ser libre, Apache Web Server tiene un gran conjunto de funcionalidades de gran alcance.
<i>Es gratuito y libre</i>	Una segunda ventaja relacionada con el diseño de código abierto de Apache es su costo. Apache es gratuito y puede ser

⁸² <http://www.cherokee-project.com/>

⁸³ <http://www.iis.ne>

⁸⁴ <http://tomcat.apache.org>

⁸⁵ <http://httpd.apache.org>

⁸⁶ <http://www.jboss.org>

⁸⁷ <http://glassfish.java.net/>

	descargado por cualquier persona.
<i>Soporte</i>	El servidor web Apache al ser muy difundido en el medio, posee una amplia gama de documentación y soporte.

Tabla 29 Ventajas del servidor web APACHE

Después de definir el servidor web, el *formulario web* está desarrollado en el lenguaje de programación **PHP**. Para mayor facilidad, se ha instalado el sistema de infraestructura **WAMP**⁸⁸.

El sistema de infraestructura **WAMP** se deriva de las siguientes herramientas:

- **Windows**, como sistema operativo, también existe alternativas para sistemas **LINUX** y **Macintosh**. Para el alcance del proyecto se instaló en un sistema operativo Windows.
- **Apache**, como servidor web. Ya se definió para alojar la aplicación web del proyecto.
- **MySQL**, como gestor de bases de datos. Para el alcance del proyecto servirá para alojar los datos del usuario que se necesiten.
- **PHP** (generalmente), **Perl**, o **Python**, como lenguajes de programación. Para el alcance del proyecto se definió **PHP** como el lenguaje de la aplicación web, por ser un lenguaje bien estructurado y muy usado en el medio.

5.4.2. Implementación de formularios web

El desarrollo de los formularios se basó en la plataforma de desarrollo web, descrita en la sección anterior, debido a que el lenguaje de implementación es **PHP**, para los alcances del proyecto se decidió crear formularios *html* con solicitud de respuesta a archivos *php* en el servidor *wamp*. Cada formulario genera una petición de almacenamiento de datos al archivo correspondiente, y el servidor procesa la solicitud.

Cuando un formulario es llenado y enviado, la aplicación web remite los datos al módulo *MIT* que se encarga de poblar la ontología del usuario con dichos datos, tal y como representa la figura **39** donde las partes en gris representan los formularios web.

⁸⁸ <http://www.wampserver.com/en/>

Las figuras **40 y 41** detallan las interfaces del prototipo para el registro de usuario y para el ingreso/edición/eliminación de preferencias explícitas respectivamente, implementadas en el servidor *WAMP*.

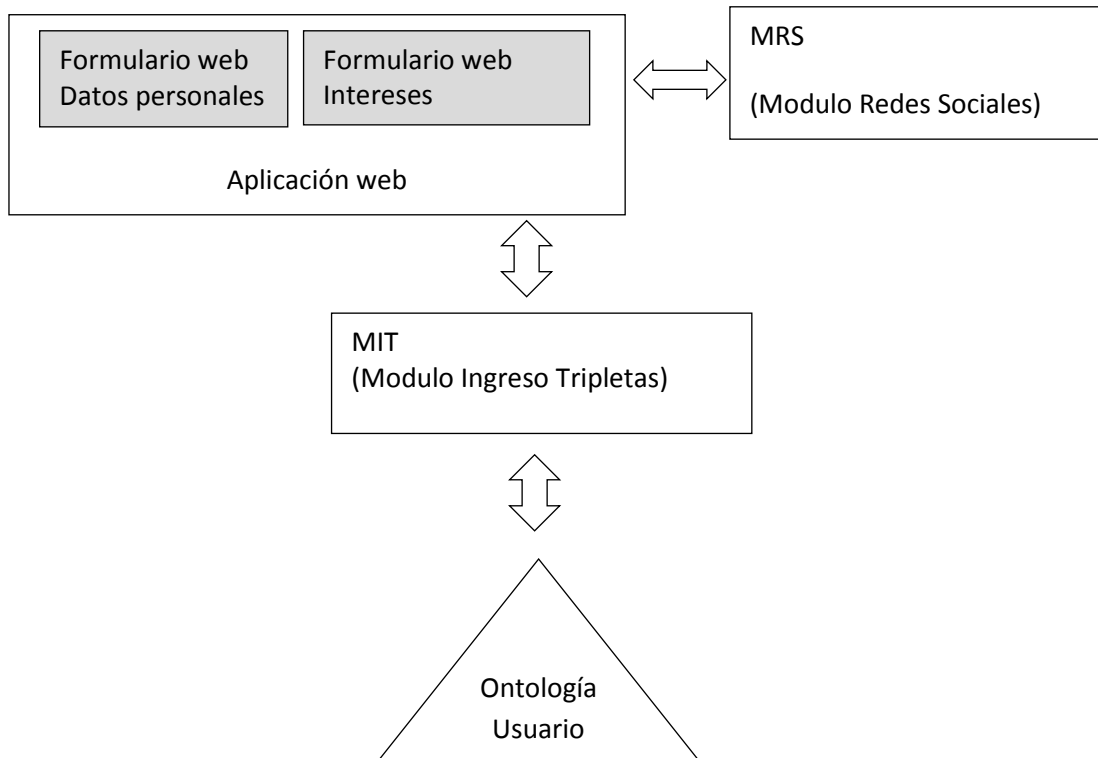
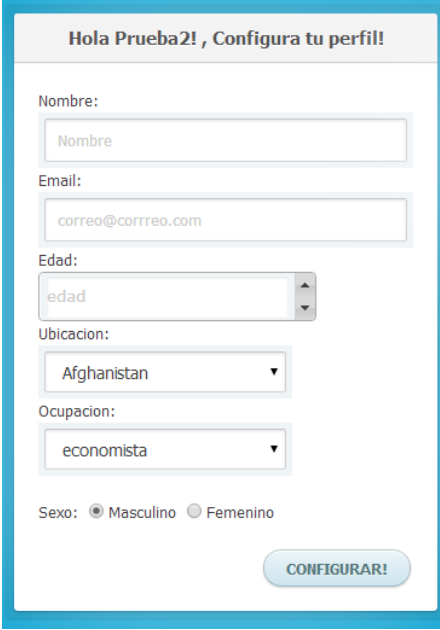


Figura 39 Formularios web integrados en la aplicación web del prototipo



Figura 40 Interfaz CRUD ⁸⁹ preferencias de usuario en el prototipo

⁸⁹ Create, Read, Update and Delete (Creación, lectura, actualización y eliminación)



The form is titled "Hola Prueba2! , Configura tu perfil!". It contains the following fields:

- Nombre: A text input field with the placeholder "Nombre".
- Email: A text input field with the placeholder "correo@correo.com".
- Edad: A dropdown menu with the placeholder "edad".
- Ubicacion: A dropdown menu with "Afghanistan" selected.
- Ocupacion: A dropdown menu with "economista" selected.
- Sexo: Two radio buttons labeled "Masculino" (selected) and "Femenino".
- A "CONFIGURAR!" button at the bottom right.

Figura 41 Interfaz formulario web para datos personales de usuario en el prototipo

5.4.3. Implementación de Redes sociales

El módulo de redes sociales MRS, es un módulo independiente desarrollado bajo la misma plataforma web. En el prototipo la aplicación web solicita a través de servicios REST los datos del usuario registrado en el prototipo, si el usuario no se encuentra configurado con una red social, el módulo MRS lo configura utilizando los métodos explicados en el capítulo 4. La implementación de este módulo en el prototipo esta descrito en la figura **44** donde la parte gris representa el modulo **MRS**.

Las figuras **42 y 43** detallan las interfaces del prototipo para consumir los datos desde las redes sociales.



Figura 42 Interfaz de registro de datos con redes sociales



Figura 43 Interfaz de solicitud de permisos de Facebook y Google+.

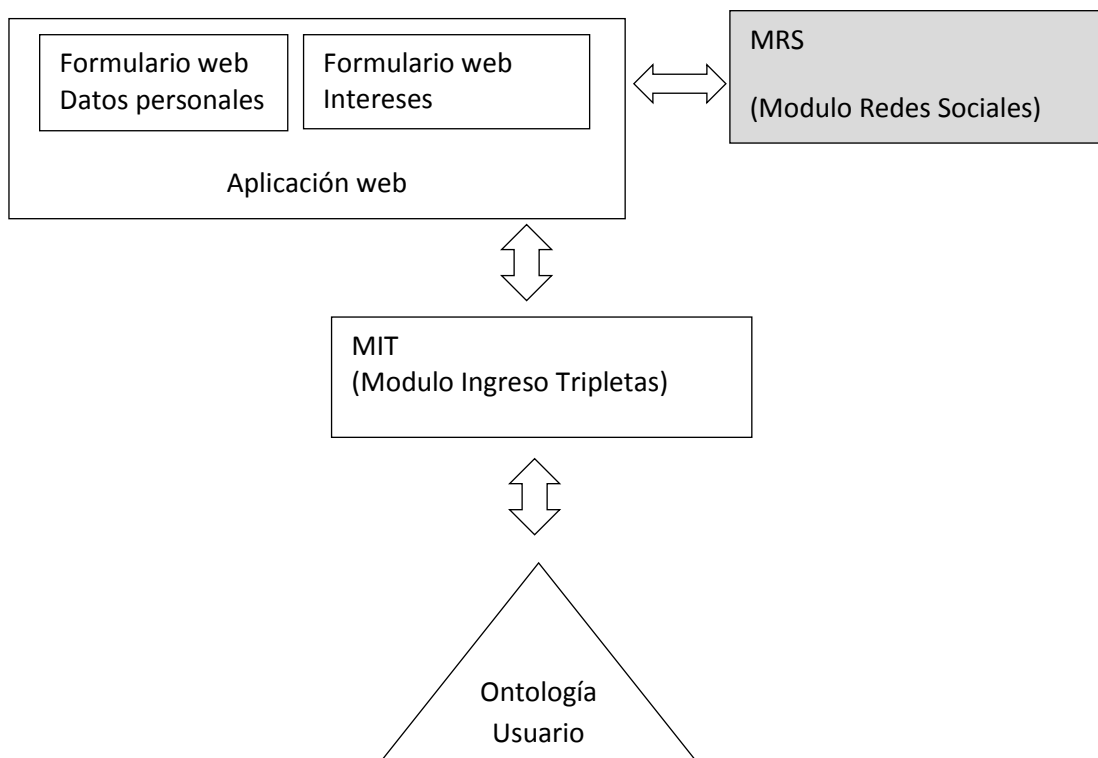


Figura 44 Modulo MRS definido en la estructura del prototipo

5.4.4. Implementación de la Aplicación Web

Después de definir el desarrollo de los formularios web y el módulo **MRS** queda por definir la aplicación web que comprende: autenticación de los usuarios, el registro y el perfil principal de cada usuario registrado. La primera interfaz de usuario, es la de inicio de sesión **al prototipo**. Esta interfaz permite escoger

entre autenticarse al sistema o registrarse en la aplicación.⁹⁰ La figura 45 muestra la interfaz de inicio de sesión a la aplicación prototipo.

The image shows a web form titled "Login to TV DIgital". It contains two input fields: "Nick" and "Password". Below these fields is a checkbox labeled "Remember me on this computer". To the right of the checkbox is a "Login" button. Below the checkbox is a "Registrate" button. The entire form is enclosed in a blue border.

Figura 45 Interfaz de inicio de sesión en el prototipo

Si el usuario decide registrarse, como se explicó en la implementación del prototipo, se muestra la pantalla de registro de usuario. Como se especificó en el capítulo 4, la fuente de información explícita, el **formulario web**, servirá para poblar de tripletas la ontología definida en el capítulo 3 de manera directa por el usuario. Las figuras 45 y 46 detallan las interfaces del prototipo para el registro de usuario y para el ingreso/edición/eliminación de preferencias explícitas respectivamente, implementadas en el servidor *WAMP*

⁹⁰ Para crear un usuario nuevo en la aplicación, se utiliza una tabla básica, que guarde *usuario* y *contraseña*



The image shows a user registration form titled "Hola Prueba2!, Configura tu perfil!". The form is divided into two main sections. The left section contains input fields for "Nombre:" (with a placeholder "Nombre"), "Email:" (with a placeholder "correo@correo.com"), "Edad:" (a dropdown menu with "edad" selected), "Ubicacion:" (a dropdown menu with "Afghanistan" selected), "Ocupacion:" (a dropdown menu with "economista" selected), and "Sexo:" with radio buttons for "Masculino" (selected) and "Femenino". A "CONFIGURAR!" button is at the bottom of this section. The right section is titled "Configurar con una red social" and features three circular icons for Facebook, Twitter, and Google+. Below these icons is a large blue rectangular area.

Figura 46 Interfaz de registro de usuario en el prototipo

Si el usuario ya está registrado en la aplicación prototipo, se procede a mostrar su **perfil** que se componen de sus datos personales y sus preferencias (si es que las tuviera). En la parte derecha de la aplicación web, se mostrara los datos personales, y las redes sociales para que pueda activarlas si lo desea, y en la parte central tendrá sus preferencias para modificarlas a su gusto, la figura **47** muestra la ubicación de cada funcionalidad descrita en la interfaz del perfil de usuario en la aplicación web prototipo.

5.4.5. Implementación de la recolección de datos desde la TV.

Para el caso de la recuperación de los datos del usuario desde la televisión digital, hay que tener en cuenta que **el usuario debió registrarse previamente y haber tenido “actividad” en la televisión digital.**

Si es así, al presionar el botón “Obtener Interacción TV Digital” (figura **47**) el prototipo recupera los datos de la televisión digital del usuario autenticado, accediendo a la base de datos, y almacenando esta información en el repositorio mediante el módulo *MIT*.



UNIVERSIDAD DE CUENCA

Hola Usuario3! Bienvenido a tu perfil de TVDIGITAL!

Salir

Obtener Interaccion Tv Digital

Tus Datos personales

age: 24
based_near: Cuenca, Ecuador
gender: male
mbox: abuelcore@hotmail.com
name: Sebastian Roman
title: Computer Science

Actualizar Datos Personales

Actualizar Tus Redes Sociales



Actualizar!



Actualizar!



Actualizar!

Intereses

Consulta

Agregar

Editar

Eliminar

Intereses

URI	Interes	Calificacion	Categoria
http://ucuenca.edu.ec/userprofile/Facebook11177446882	Kevin Spacey	3	Actor/direct
http://ucuenca.edu.ec/userprofile/Facebook8146878809	Hypocrisy	3	Musician/ba
http://ucuenca.edu.ec/userprofile/Facebook363184403788142	Profidental	3	Professiona
http://ucuenca.edu.ec/userprofile/Facebook126844294034258	John Petrucci	3	Musician/ba
http://ucuenca.edu.ec/userprofile/Facebook453679927986630	Ecuador Es Un Pais de Ripley	3	Entertainme
http://ucuenca.edu.ec/userprofile/Facebook379827575396081	La Patada Giratoria de Chuck Norris	3	Cause

⌂ ⌕

Página 1 de 23 10

Mostrando 1 - 9 de 228

Figura 47 Interfaz de perfil de usuario en el prototipo

5.5. PRUEBAS Y RESULTADOS

5.5.1. Fundamentos de la prueba

Cuando se desarrolla un prototipo de software existen posibilidades de que se produzcan fallos humanos en las actividades de producción. Pueden darse errores en cualquier fase del proceso, desde una mala especificación de los objetivos, hasta errores en las fases de diseño y desarrollo.

Es imposible que el ser humano trabaje de una manera perfecta, y por ello, el desarrollo debe ir acompañado de una actividad que garantice la calidad. Así pues, la prueba del prototipo de software y la ontología del usuario representa la revisión final del proyecto.

La fase de pruebas del proyecto tiene dos alcances: **Pruebas del prototipo** desarrollado en ambiente web, y las **Pruebas de la ontología** que modela el usuario de la televisión digital. Dentro de estas fases se pueden desarrollar varios tipos distintos de pruebas en función de los objetivos de las mismas.

5.5.1.1. Pruebas del prototipo

Uno de los objetivos de la fase de pruebas del proyecto es verificar que el comportamiento externo del prototipo satisfaga los requisitos previamente definidos en la fase de análisis del prototipo.

Una de las técnicas más empleadas para la especificación funcional de prototipos de software son los casos de uso en ejecución. A continuación se presentarán los requisitos del prototipo junto con su caso de uso.

- Sistema de autenticación de usuario.

Fundamentos: El usuario accede a la aplicación y desea autenticarse. Ingresa sus datos y pulsa “login”.

Caso de prueba 1: El usuario accede correctamente, se muestra su respectivo perfil en el prototipo. Ver figura 48.

Caso de prueba 2: El usuario desea acceder con datos erróneos, se muestra mensaje de error. Ver figura 49.

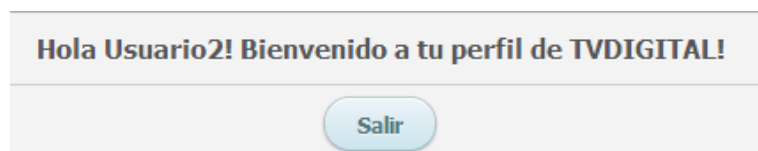
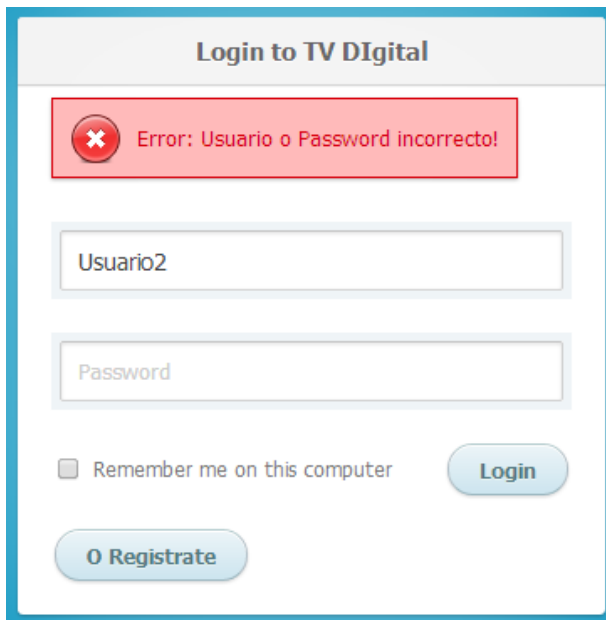


Figura 48 Login correcto, saludo inicial del perfil de usuario en el prototipo



The screenshot shows a web form titled "Login to TV Digital". At the top, there is a red error message box with a white 'x' icon and the text "Error: Usuario o Password incorrecto!". Below this, there are two input fields: the first is labeled "Usuario2" and the second is labeled "Password". Under the "Password" field, there is a checkbox labeled "Remember me on this computer". To the right of the checkbox is a blue "Login" button. Below the checkbox is a blue button labeled "O Registrate".

Figura 49 Login incorrecto, mensaje de error al usuario

Resultado: El sistema muestra en cada caso la interfaz correspondiente al usuario.

Conclusión: Prueba superada con éxito. Se muestra en cada caso la interfaz correspondiente, como se puede comprobar en las figuras **48 y 49**.

- Registro de usuario desde un formulario en la web.

Fundamentos: El usuario selecciona registrarse en el prototipo, ingresa sus datos personales al formulario web y pulsa "*configurar*".

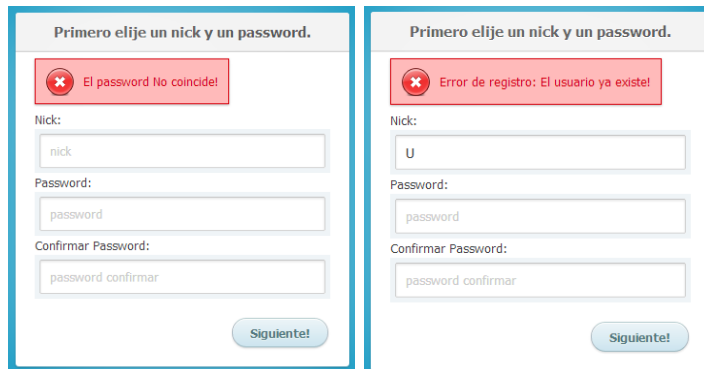
Caso de prueba 1: El prototipo solicita al usuario su identificador y contraseña respectiva para configurarlo en la base de datos. El usuario ingresa correctamente, se muestra la respectiva interfaz para el registro con formulario en el prototipo. Ver figura **50**.

Caso de prueba 2: El prototipo solicita al usuario su identificador y contraseña respectiva para configurarlo en la base de datos. El usuario ingresa erróneamente la contraseña de confirmación, o selecciona un identificador ya utilizado. Se muestra el mensaje de error respectivo al usuario. Ver figura **51**.

Caso de prueba 3: El usuario accede al formulario web, el prototipo solicita al usuario los datos. El usuario ingresa correctamente los datos en el formulario y

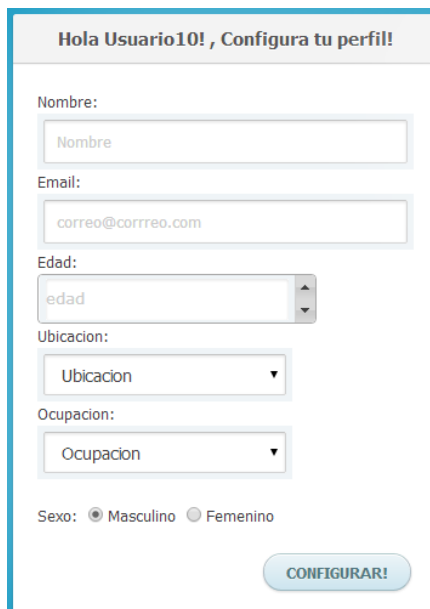
pulsa “configurar”. El prototipo almacena los datos del usuario. El prototipo muestra la interfaz de autenticación con un mensaje de que ha sido configurado correctamente el usuario. Ver figura 53.

Caso de prueba 4: El usuario accede al formulario web, el prototipo solicita al usuario los datos. El usuario no ingresa los campos obligatorios en el formulario y pulsa “configurar”. El prototipo solicita que se ingrese los campos obligatorios. Ver figura 53.



The figure shows two side-by-side screenshots of a web form titled "Primero elije un nick y un password." The form has three input fields: "Nick:", "Password:", and "Confirmar Password:". Below the fields is a "Siguiente!" button. In the left screenshot, a red error message box at the top says "El password No coincide!". In the right screenshot, a red error message box at the top says "Error de registro: El usuario ya existe!".

Figura 50 Mensajes de error en la creación del usuario



The figure shows a single screenshot of a web form titled "Hola Usuario10! , Configura tu perfil!". The form contains several input fields and a "CONFIGURAR!" button. The fields are: "Nombre:" (text input), "Email:" (text input with placeholder "correo@correo.com"), "Edad:" (text input with a dropdown arrow), "Ubicacion:" (text input with a dropdown arrow), "Ocupacion:" (text input with a dropdown arrow), and "Sexo:" (radio buttons for "Masculino" and "Femenino").

Figura 51 Interfaz de registro de datos personales con formulario web.



Figura 52 Mensaje de solicitud para llenar campos obligatorios

Figura 53 Mensaje de registro correcto del usuario.

Resultado: El sistema muestra en cada caso la interfaz correspondiente al usuario.

Conclusión: Prueba superada con éxito. Se muestra en cada caso la interfaz correspondiente, como se puede comprobar en las figuras **52 y 53**.

- Registro de usuario accediendo a sus redes sociales.

Fundamentos: El usuario selecciona registrarse en el prototipo y accede a los datos de sus redes sociales pulsando cualquier red social "Facebook". "Twitter" o "Google".

Caso de prueba 1: El prototipo solicita al usuario su identificador y contraseña respectiva para configurarlo en la base de datos. El usuario ingresa correctamente, se muestra la respectiva interfaz para el registro con formulario en el prototipo. Ver figura **45**.

Caso de prueba 2: El prototipo solicita al usuario su identificador y contraseña respectiva para configurarlo en la base de datos. El usuario ingresa erróneamente la contraseña de confirmación, se muestra el mensaje respectivo al usuario. Ver figura **50**.

Caso de prueba 3: El usuario accede a su red social, pulsando el botón de “Facebook”. “Twitter” o “Google”. El prototipo redirecciona al usuario al módulo para la configuración de la red social. El módulo redirige al usuario a la página de ingreso de la red social seleccionada (Ver figura **54** y **55**), el usuario ingresa sus datos y otorga los permisos respectivos. El prototipo almacena los datos del usuario devueltos por el modulo **MRS**. El modulo muestra el respectivo mensaje. Ver figura **56** y **57**.

Caso de prueba 4: El usuario accede a su red social, pulsando el botón de “Facebook”. “Twitter” o “Google”. El prototipo redirecciona al usuario al módulo para la configuración de la red social. El módulo redirige al usuario a la página de ingreso de la red social seleccionada, el usuario decide **no** ingresar sus datos y **no** otorga los permisos respectivos. El prototipo no almacena los datos del usuario. El modulo muestra un mensaje de error. Ver figura **58**.



Figura 54 Interfaz de autenticación con la red social Facebook.

¿Autorizas a PruebaV para que utilice tu cuenta?

Esta aplicación **será capaz de:**

- Leer Tweets de tu cronología.
- Ver a quién sigues.

☐ Recordar mis datos · ¿Olvidaste tu contraseña?

Inicia Sesión **Cancelar**

Esta aplicación **no tendrá capacidad para:**

- Seguir a nuevas personas.
- Actualizar tu perfil.
- Publicar Tweets por ti.
- Acceder a tus mensajes directos.
- Ver tu contraseña de Twitter.

Acceder con la cuenta de Google



Acceder

☒ No salir de la cuenta · [¿Necesitas ayuda?](#)

Figura 55 Interfaz de autenticación con las redes sociales Twitter y Google+.

Felicitaciones! Tu usuario hybridauth fue configurado exitosamente: Usuario12, con el proveedor: Twitter. Y ya esta listo para usar!

Por favor regresa a tu aplicacion : [Clic aqui](#)

Figura 56 Mensaje de registro correcto en el módulo MRS.




TvDigitalPrueba2 recibirá la siguiente información tuya: **perfil público**, lista de amigos, dirección de correo electrónico, listas de amigos personalizadas, Últimas noticias, cumpleaños, historial académico, localidad natal, ciudad actual, sitio web, descripción personal y clics en "Me gusta".

 No se permite que la aplicación publique en Facebook.

Cancelar **Aceptar**

Esta aplicación solicita los siguientes permisos:

-  Conocer tu nombre, información básica y la lista de personas con las que estás conectado en Google+ (Editar lista)
-  Ver tu dirección de correo electrónico
-  Permitir que Google comunique a las personas de estos círculos que has iniciado sesión en esta aplicación con Google:

 **Tus círculos**

+ Añadir a más personas
- ☐ Solo tú

PruebaSDKProgrammingApps y Google usarán esta información de acuerdo con sus respectivas condiciones de servicio y políticas de privacidad.

Cancelar **Aceptar**

Figura 57 Interfaz de solicitud de permisos de Facebook y Google+.

Details

Type: Exception
Code: 5
Message: Authentication failed! The user denied your request.
File: C:\wamp\www\MRS\hybridauth\Hybrid\Auth.php
Line: 149

Figura 58 Mensaje de registro incorrecto en el módulo MRS.

Resultado: El sistema muestra en cada caso la interfaz correspondiente al usuario.

Conclusión: Prueba superada con éxito. Se muestra en cada caso la interfaz correspondiente, como se puede comprobar en las figuras **57 y 58**.

- Registro de usuario desde un formulario de ingreso en la web y accediendo a sus redes sociales.

Fundamentos: El usuario selecciona registrarse en el prototipo con el formulario web y después accede a los datos de sus redes sociales pulsando cualquier red social “Facebook”. “Twitter” o “Google”.

Caso de prueba 1: El prototipo solicita al usuario su identificador y contraseña respectiva para configurarlo en la base de datos. El usuario ingresa correctamente, se muestra la respectiva interfaz para el registro con formulario en el prototipo. Ver figura **45**.

Caso de prueba 2: El prototipo solicita al usuario su identificador y contraseña respectiva para configurarlo en la base de datos. El usuario ingresa erróneamente la contraseña de confirmación, se muestra el mensaje respectivo al usuario. Ver figura **50**.

Caso de prueba 3: El usuario accede al formulario web, el prototipo solicita al usuario los datos. El usuario ingresa correctamente los datos en el formulario y pulsa “configurar”. El prototipo almacena los datos del usuario. El prototipo muestra la interfaz de autenticación con un mensaje de que ha sido configurado correctamente el usuario. Ver figura **51**.

Caso de prueba 4: El usuario accede al formulario web, el prototipo solicita al usuario los datos. El usuario no ingresa los campos obligatorios en el formulario y pulsa “configurar”. El prototipo solicita que se ingrese los campos obligatorios. Ver figura **52**.

Caso de prueba 5: El usuario accede a su red social dentro del perfil, pulsando el botón de “Facebook”. “Twitter” o “Google”. Si es la primera vez que accede con la red social seleccionada, el prototipo redirecciona al usuario al módulo para la configuración de la red social. El módulo dirige al usuario a la página de ingreso de la red social seleccionada, el usuario ingresa sus datos y otorga los permisos respectivos. El prototipo almacena los datos del usuario devueltos por el modulo **MRS**. El modulo muestra el respectivo mensaje. Ver figura **53**.

Caso de prueba 6: El usuario accede a su red social dentro de su perfil, pulsando el botón de “Facebook”. “Twitter” o “Google”. El prototipo redirecciona al usuario al módulo para la configuración de la red social. El módulo dirige al usuario a la página de ingreso de la red social seleccionada, el usuario decide **no** ingresar sus datos y **no** otorga los permisos respectivos. El prototipo no almacena los datos del usuario. El módulo muestra un mensaje de error. Ver figura **54**.

Resultado: El sistema muestra en cada caso la interfaz correspondiente al usuario.

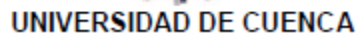
Conclusión: Prueba superada con éxito. Se muestra en cada caso la interfaz correspondiente, como se puede comprobar en las figuras **57 y 58**.

- Manejo de preferencias de un usuario.

Fundamentos: El usuario selecciona modificar sus preferencias en el prototipo con el formulario web.

Caso de prueba 1: El usuario ingresa en el prototipo y selecciona agregar un interés personal, el prototipo presenta el formulario de agregar preferencias, el usuario ingresa correctamente los datos, el prototipo presenta mensaje de que ha sido agregado correctamente. Ver figura **59**.

Caso de prueba 2: El usuario ingresa en el prototipo y selecciona editar un interés seleccionado, el prototipo presenta el formulario de editar preferencias, el



Caso de prueba 3: El usuario accede al formulario web, el prototipo solicita al usuario los datos. El usuario ingresa correctamente los datos en el formulario y pulsa “configurar”. El prototipo almacena los datos del usuario. El prototipo muestra la interfaz de autenticación con un mensaje de que ha sido configurado correctamente el usuario. Ver figura **61**.

Caso de prueba 4: El usuario ingresa en el prototipo y selecciona eliminar un interés seleccionado, el prototipo presenta un mensaje de confirmación para eliminar la preferencia, el usuario acepta, el prototipo presenta mensaje de que ha sido eliminado correctamente. Ver figura **62**.

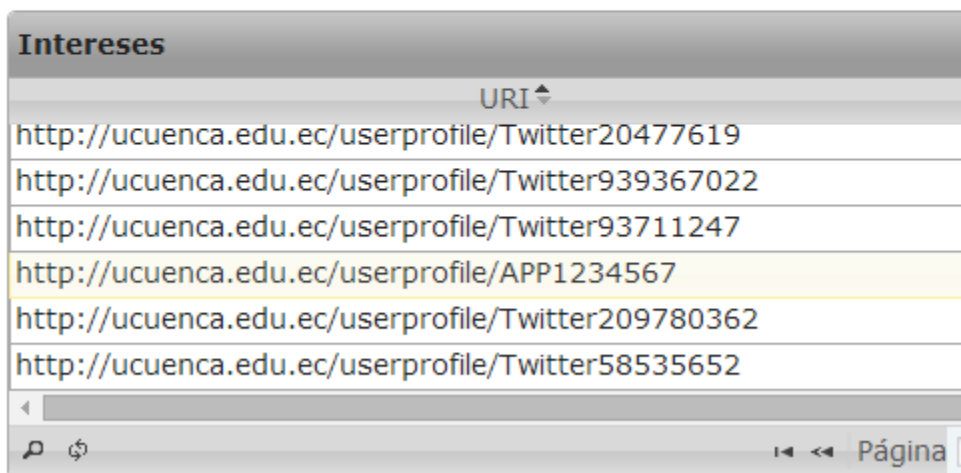


Figura 59 Interfaz para el CRUD de preferencias.

Agregar nuevo Interes x

All form fields are required.

Id

1234567

Interes

Interes

Doi

5

Agregar

Cancelar

Mensaje de la página localhost:

Agregado exitoso!

Aceptar

Figura 60 Interfaz para agregar nuevas preferencias

Editar Interes x

All form fields are required.

Id

http://ucuenca.edu.ec/userp

Interes

Interes

Doi

4

Editar

Cancelar

Mensaje de la página localhost:

Editado exitoso!

Aceptar

Figura 61 Interfaz para editar preferencias.



Figura 62 Mensaje para eliminar preferencias

Resultado: El sistema muestra en cada caso la interfaz correspondiente al usuario.

Conclusión: Prueba superada con éxito. Se muestra en cada caso la interfaz correspondiente, como se puede comprobar en las figuras **61 y 62**.

5.5.1.2. Pruebas de Consistencia en la Ontología

Después de las pruebas del prototipo para la recolección de datos del usuario mediante las técnicas implícitas y explícitas definidas, se procede a generar las pruebas sobre la ontología definida en el capítulo 3.

Para el alcance del proyecto, se ha decidido probar el modelo ontológico, con consultas **SPARQL**⁹¹, que es lenguaje de consultas diseñado para repositorios semánticos. Las consultas que serán tomadas para las pruebas son las consultas definidas en la sección 3.2 en la tabla 5, *preguntas personales*.

Para poder realizar las pruebas de la ontología, se creó una interfaz en el propio perfil de usuario, donde soporte consultas **sparql** generadas, y genere respuestas directo en la aplicación web. La figura 63 muestra la interfaz descrita que permite realizar consultas **sparql** para poder testar el modelo de usuario para la televisión digital especificado en el capítulo 3.

⁹¹ <http://www.w3.org/TR/rdf-sparql-query>



UNIVERSIDAD DE CUENCA

Intereses Consulta

Consulta 1 ▼

CARGAR CONSULTA!

Consulta: Consulta Sparql

CONSULTAR!

Figura 63 Interfaz para realizar consultas sparql en el prototipo.

A continuación se procede a responder preguntas seleccionadas de la tabla **5** mediante consultas **Sparql** en la interfaz del prototipo.

- PC1. ¿Cuál es el nombre del televidente?

Fundamentos: El usuario desea saber el nombre del televidente. Ver figura **64**.

Consulta Sparql:

```
1 select    ?name    from    <http://ucuenca.edu.ec/userprofile/>    where{
2 <http://ucuenca.edu.ec/userprofile/usuario>    <http://xmlns.com/foaf/0.1/name>
3 ?name}
4
```



UNIVERSIDAD DE CUENCA

Seleccionar una consulta:

Cual es el nombre del televidente? ▼

CARGAR CONSULTA!

Consulta:

```
select ?name from <http://ucuenca.edu.ec/userprofile/> where
{
  <http://ucuenca.edu.ec/userprofile/Usuario2> <http://xmlns.com/foaf/0.1/name> ?name}
```

CONSULTAR!

RESPUESTA	
name	
Danielita Roman Jarrin	

Figura 64 Prueba PC1. Respuesta correcta del nombre del televidente.

Resultado: El sistema muestra el correspondiente nombre del televidente.

Conclusión: Prueba superada con éxito.

- PC2. ¿Cuál es la ciudad de residencia del televidente?

Fundamentos: El usuario desea saber el lugar donde vive un televidente en específico. El tipo de residencia de la ontología viene dado por ciudad-país, dependiendo el método de recolección de datos. Ver figura 65.

Consulta Sparql:

```
1 select ?residencia from <http://ucuenca.edu.ec/userprofile/> where
2     {
3       <http://ucuenca.edu.ec/userprofile/usuario>
4       <http://xmlns.com/foaf/0.1/name> ?residencia}
```

Seleccionar una consulta:

Cual es la ciudad de residencia del te ▼

CARGAR CONSULTA!

Consulta: `select ?residencia from <http://ucuenca.edu.ec/userprofile/> where
{
<http://ucuenca.edu.ec/userprofile/Usuario3> <http://xmlns.com/foaf/0.1/based_near> ?residencia }`

CONSULTAR!

RESPUESTA

residencia

Cuenca, Ecuador

Figura 65 Prueba PC2. Respuesta correcta de la residencia del televidente.

Resultado: El sistema muestra la correspondiente residencia del televidente.

Conclusión: Prueba superada con éxito

- PC3. ¿Cuál es el rango de edad del televidente?

Fundamentos: El usuario desea saber la edad del televidente. Con la edad el televidente pueda definir el rango: joven, adulto, etc. Ver figura 66.

Consulta Sparql:

```
1 select ?edad from <http://ucuenca.edu.ec/userprofile/> where
2 {
3   <http://ucuenca.edu.ec/userprofile/usuario>
4   <http://xmlns.com/foaf/0.1/age> ?edad}
```



UNIVERSIDAD DE CUENCA

Seleccionar una consulta:

Cual es el rango de edad del televidente ▼

CARGAR CONSULTA!

Consulta:

```
select ?edad from <http://ucuenca.edu.ec/userprofile/> where
{
  <http://ucuenca.edu.ec/userprofile/Usuario3> <http://xmlns.com/foaf/0.1/age> ?edad}
```

CONSULTAR!

RESPUESTA

edad

24

Figura 66 Prueba PC3. Respuesta correcta de la edad del televidente.

Resultado: El sistema muestra la correspondiente edad del televidente.

Conclusión: Prueba superada con éxito.

- PC4. ¿Cuáles son los tópicos de interés del televidente?

Fundamentos: El usuario desea saber las categorías de las preferencias del televidente. Ver figura 67.

Consulta Sparql:

```
1 select DISTINCT ?Categoria from <http://ucuenca.edu.ec/userprofile/> where
2     {
3         <http://ucuenca.edu.ec/userprofile/usuario>
4         <http://xmlns.com/foaf/0.1/interest> ?o.
5         ?o <http://xmlns.com/foaf/0.1/interest_category> ?Categoria}
6
```


Seleccionar una consulta:

Cuales son los topicos de interes de ▾

CARGAR CONSULTA!

Consulta: `select DISTINCT ?Categoria from <http://ucuenca.edu.ec/userprofile/> where { <http://ucuenca.edu.ec/userprofile/Usuario3> <http://xmlns.com/foaf/0.1/interest> ?o . ?o <http://xmlns.com/foaf/0.1/interest_category> ?Categoria }`

CONSULTAR!

RESPUESTA	
Categoria	
Website	
Actor/director	
Musician/band	
Professional services	
Entertainment website	
Cause	
Video game	
City	
Movie	
default	
Community	
Non-profit organization	
Professional sports team	

Figura 67 Prueba PC4. Respuesta correcta sobre los tópicos de interés del televidente.

Resultado: El sistema muestra todas las categorías correspondientes a las preferencias del televidente.

Conclusión: Prueba superada con éxito

- PC5. ¿Qué ve el televidente en las mañanas-tardes-noches?

Fundamentos: El usuario desea saber los programas de televisión que mira en determinada hora. Con la hora se puede definir mañana, tarde, y noche. Ver figura 68.

Consulta Sparql:

```
1 select ?Programa from <http://ucuenca.edu.ec/userprofile/> where
2 {
3   <http://ucuenca.edu.ec/userprofile/usuario> <http://xmlns.com/foaf/0.1/interaccion>?o.
4   ? o <http://xmlns.com/foaf/0.1/Programme> ?Programa .
5   ?o <http://xmlns.com/foaf/0.1/schedule_date> ?p FILTER regex(?p, "11:00")}
```

6
7

Seleccionar una consulta:

Que ve el televidente en las mañanas ▾

CARGAR CONSULTA!

Consulta:

```
select ?Programa from <http://ucuenca.edu.ec/userprofile/> where
{
  <http://ucuenca.edu.ec/userprofile/Usuario3> <http://xmlns.com/foaf/0.1/interaccion> ?o . ?o <http://xmlns.com/foaf/0.1/Programme> ?
  Programa . ?o <http://xmlns.com/foaf/0.1/schedule_date> ?p FILTER regex(?p, "11:00")}
```

CONSULTAR!

RESPUESTA

Programa

El precio de la historia
The Boys

Figura 68 Prueba PC5. Respuesta correcta de programas que mira a las 11:00h el televidente.

Resultado: El sistema muestra todos los programas vistos a la hora correspondiente del televidente.

Conclusión: Prueba superada con éxito

- PC6. ¿Qué ve el televidente los fines de semana?

Fundamentos: El usuario desea saber los programas de televisión que mira en determinado día. Con el día se puede determinar los fines de semana. Ver figura 69.

Consulta Sparql:

```
1 select ?Programa from <http://ucuenca.edu.ec/userprofile/> where
2 {
3   <http://ucuenca.edu.ec/userprofile/usuario> <http://xmlns.com/foaf/0.1/interaccion> ?o
4   .
5   ?o <http://xmlns.com/foaf/0.1/Programme> ?Programa .
6   ?o <http://xmlns.com/foaf/0.1/schedule_date> ?p FILTER regex(?p, "2014-08-06")}
```

Seleccionar una consulta:

Que ve el televidente los fines de se ▾

CARGAR CONSULTA!

Consulta: `select ?Programa from <http://ucuenca.edu.ec/userprofile/> where { <http://ucuenca.edu.ec/userprofile/Usuario3> <http://xmlns.com/foaf/0.1/interaccion> ?o . ?o <http://xmlns.com/foaf/0.1/Programme> ?Programa . ?o <http://xmlns.com/foaf/0.1/schedule_date> ?p FILTER regex(?p, "2014-08-06") }`

CONSULTAR!

RESPUESTA	
Programa	
Noticias	
ESPN Report	

Figura 69 Prueba PC6. Respuesta correcta de programas que mira el día domingo “2014-08-06” el televidente.

Resultado: El sistema muestra todos los programas vistos el día correspondiente.

Conclusión: Prueba superada con éxito.

- PC7. ¿Cuál es su personaje favorito?

Fundamentos: El usuario desea saber los personajes favoritos de televisión del televidente. Para ello se busca todos los intereses que tengan la categoría “Actor”. Ver figura 70.

Consulta Sparql:

```

1  select DISTINCT ?Actor from <http://ucuenca.edu.ec/userprofile/> where
2      {
3      <http://ucuenca.edu.ec/userprofile/usuario> <http://xmlns.com/foaf/0.1/interest>?o.
4      ? o <http://xmlns.com/foaf/0.1/interest_name> ?Actor.
5      ?o <http://xmlns.com/foaf/0.1/interest_category> ?a FILTER regex(?a, "Actor")
6
7
8

```

Seleccionar una consulta:

Cual es su personaje favorito?

CARGAR CONSULTA!

Consulta:

```
select DISTINCT ?Actor from <http://ucuenca.edu.ec/userprofile/> where
{
  <http://ucuenca.edu.ec/userprofile/Usuario3> <http://xmlns.com/foaf/0.1/interest> ?o . ?o <http://xmlns.com/foaf/0.1/interest_name> ?
  Actor. ?o <http://xmlns.com/foaf/0.1/interest_category> ?a FILTER regex(?a, "Actor")}
```

CONSULTAR!

RESPUESTA

Actor

Kevin Spacey
Jennifer Lawrence
Cuenca Pedacito De CIELO
Maisie Williams
Morgan Freeman
Alfred Hitchcock
Quentin Tarantino
Edward Norton
Bill Murray
Maisie_Williams
Emilia Clarke

Figura 70 Prueba PC7. Respuesta correcta de los personajes favoritos del televidente

Resultado: El sistema muestra todos los personajes favoritos del televidente.

Conclusión: Prueba superada con éxito.

- PC8. ¿Cuál es la profesión del televidente?

Fundamentos: El usuario desea saber la profesión del televidente. Ver figura 71.

Consulta Sparql:

```
1 select ?profesion from <http://ucuenca.edu.ec/userprofile/> where
2 {
3   <http://ucuenca.edu.ec/userprofile/usuario> <http://xmlns.com/foaf/0.1/title>
4   ?profesion}
5
```

Seleccionar una consulta:

Cual es la profesion del televidente? ▾

CARGAR CONSULTA!

Consulta: `select ?profesion from <http://ucuenca.edu.ec/userprofile/> where
{
<http://ucuenca.edu.ec/userprofile/Usuario3> <http://xmlns.com/foaf/0.1/title> ?profesion}`

CONSULTAR!

RESPUESTA	
profesion	
Computer Science	

Figura 71 Prueba PC8. Respuesta correcta de la profesión del televidente.

Resultado: El sistema muestra la correspondiente profesión del televidente.

Conclusión: Prueba superada con éxito



UNIVERSIDAD DE CUENCA

CONCLUSIONES Y TRABAJOS FUTUROS

A continuación se presentarán las conclusiones del trabajo realizado y se propondrán trabajos futuros que complementen a esta tesis.

CAPÍTULO 6

CONCLUSIONES Y TRABAJOS FUTUROS

En el capítulo actual se presentan las conclusiones y recomendaciones de la tesis “MODELAMIENTO DEL PERFIL DE UN USUARIO USANDO TECNOLOGIAS SEMANTICAS” en base a los objetivos planteados en el capítulo 1. Primero se incluyen las conclusiones de cada uno de los objetivos, posteriormente recomendaciones y trabajos futuros que puedan aportar al proyecto de televisión digital del cual es parte esta tesis.

6.1. CONCLUSIONES

Como observamos en el desarrollo del proyecto, el modelamiento de un perfil de usuario puede tomar 2 aspectos principales: su perfil estático que comprende la información principal de un usuario y el perfil dinámico que comprende sus intereses y actividades. Una vez categorizados estos aspectos se pudo modelar el perfil de un usuario de televisión digital.

Después de definir los aspectos y datos de modelado de un perfil de usuario de televisión digital, considerando la metodología NeOn de modelamiento de ontologías, se concluye que FOAF es la ontología que cumple con la mayor cantidad de objetivos del perfil para el proyecto. Como siguiente paso se procede a extender dicha ontología para cumplir con la totalidad de las necesidades del modelo de perfil de usuario para televisión digital.

Una vez definido el modelo de datos ontológicos dentro del proyecto, se procede a buscar las fuentes de información que completen este perfil. De esta manera se toma como base un perfil estático llenado por el usuario mediante un formulario web de registro para posteriormente completarlo mediante el acceso a las redes sociales Facebook, Twitter y Google+. De esta manera la información obtenida cumple con los aspectos definidos en el modelo ontológico.

Mediante tecnologías semánticas se puede definir un modelo de datos que una máquina pueda interpretar gracias a las ontologías, las cuales manejan lenguaje natural de una persona por lo que la información también puede ser interpretada por un

ser humano. Gracias al modelo ontológico definido la información almacenada será accesible por humanos y máquinas dentro de un mismo modelo.

Con la construcción del prototipo de modelamiento de usuario se puede determinar que el modelo de datos obtenido es óptimo para la recolección de datos del perfil, inferencia y para consultas de datos dentro del mismo.

Como reflexión final, se concluye que los métodos de modelamiento, recolección y almacenamiento de información de perfil de usuario de televisión digital mediante tecnologías semánticas ayudan a capturar la información del usuario tanto en un arranque en frío como en un escenario ideal de datos obtenidos.

6.2. TRABAJOS FUTUROS

A partir de los resultados obtenidos con el prototipo de modelamiento de perfil de usuario se abre la posibilidad de extensión del prototipo con los siguientes trabajos posibles.

Si bien es cierto, los eventos de la televisión digital actualmente se obtienen desde una tabla estática de una base de datos, el uso de una aplicación intermedia entre el prototipo y la televisión digital permitiría procesar de mejor manera los datos de interacciones con el usuario.

Las redes sociales definidas en el alcance de este proyecto cumplen con los requisitos para el prototipo de perfil, sin embargo existen redes sociales que podrían complementar los datos de un perfil de usuario para definir mayores intereses. Por esto motivo la integración de más redes sociales a la aplicación puede cumplir este objetivo.

En el caso de las redes sociales integradas, el reconocimiento de palabras para publicaciones puede generar más información coherente sobre intereses de un usuario. Por este motivo se puede implementar un servicio de reconocimiento de lenguaje natural que facilite esta actividad.

La ontología ha sido probada mediante el cumplimiento de las preguntas planteadas pero para perfeccionarla se puede probar la ontología con nuevos métodos que evalúen otros aspectos para la misma, ejemplo usar agentes informáticos razonadores para evaluar ontologías.

Como pudimos observar los servicios REST proveen grandes ventajas dentro del alcance de esta aplicación, sin embargo un servicio adicional que permita el proceso



UNIVERSIDAD DE CUENCA

intermedio de mantener una sesión de usuario activa ayudaría a la facilidad de uso del módulo de redes sociales.

REFERENCIAS

- [1] Stuart E. Middleton, Nigel R. Shadbolt, and David C. De Roure. 2004. Ontological user profiling in recommender systems. *ACM Trans. Inf. Syst.* 22, 1 (January 2004), 54-88. DOI=10.1145/963770.963773 <http://doi.acm.org/10.1145/963770.963773>.
- [2] A. Kobsa, "User modeling: Recent work, prospects and hazards," in *Adaptive User Interfaces: Principles and Practice*, M. Schneider-Hufschmidt, T. K"uhme, and U. Malinowski, Eds. Amsterdam: North-Holland, 1993, pp. 111–128.
- [3] Fabrizio Sebastiani. 2002. Machine learning in automated text categorization. ***ACM Comput. Surv.*** 34, 1 (March 2002), 1-47. DOI=10.1145/505282.505283 <http://doi.acm.org/10.1145/505282.505283>
- [4] S. E. Middleton, D. C. De Roure, and N. R. Shadbolt. 2001. "Capturing knowledge of user preferences: Ontologies in recommender systems," in *Proceedings of the 1st International Conference on Knowledge Capture*, ser. K-CAP '01. New York, NY, USA: ACM, 2001, pp. 100–107.
- [5] S. Gauch, J. Chaffee, and A. Pretschner. 2003, "Ontology-based user profiles for search and browsing," *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, Special Issue on User Modeling for Web and Hypermedia Information Retrieval, 2003.
- [6] J.-Y. Kim, J.-W. Kim, and C.-S. Kim, "Ontology-based user preference modeling for enhancing interoperability in personalized services," in *Universal Access in Human-Computer Interaction. Applications and Services*, ser. Lecture Notes in Computer Science, C. Stephanidis, Ed. Springer Berlin Heidelberg, 2007, vol. 4556, pp. 903–912.
- [7] M. R. Tazari, M. Grimm, and M. Finke, "Modeling user context," in *Proceedings of the 10th International Conference on Human-Computer Interaction (HCI2003)*, Crete, Greece, 2003.
- [8] S. Lawrence, "Context in web search," *IEEE Data Engineering Bulletin*, vol. 23, no. 3, pp. 25–32, 2000.
- [9] J. Trajkova and S. Gauch, "Improving ontology based user profiles." in *RIAO*, C. Fluhr, G. Grefenstette, and W. B. Croft, Eds. CID, 2004, pp. 380–390.

- [10] V. Katifori, A. Poggi, M. Scannapieco, T. Catarci, and Y. E. Ioannidis, "Ontopim: how to rely on a personal ontology for personal information management." in Semantic Desktop Workshop, ser. CEUR Workshop Proceedings, S. Decker, J. Park, D. Quan, and L. Sauermann, Eds., vol. 175. CEUR-WS.org, 2005.
- [11] I. Cantador, P. Castells, and A. Bellogín, "An enhanced semantic layer for hybrid recommender systems: Application to news recommendation," *Int. J. Semant. Web Inf. Syst.*, vol. 7, no. 1, pp. 44–78, Jan. 2011. [Online]. Available: <http://dx.doi.org/10.4018/jswis.2011010103>
- [12] A. Sieg, B. Mobasher, and R. Burke, "Improving the effectiveness of collaborative recommendation with ontology-based user profiles," in *Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems*, ser. *HetRec '10*. New York, NY, USA: ACM, 2010, pp. 39–46.
- [13] N. P. Andersen, "Reducing cold start problem in the wikipedia recommender system," Master's thesis, Technical University of Denmark, DTU Informatics, Asmussens Alle, Building 305, DK-2800 Kgs. Lyngby, Denmark, 2011.
- [14] Neches, Robert; Fikes, Richard; Finin, Timothy; Gruber, Thomas; Patil, Ramesh; Senator, Edward & Swartout, William. 1991. "Enabling Technology for Knowledge Sharing". En: *Artificial Intelligence Magazine*. Vol. 12, num. 3, American Association for Artificial Intelligence Press. 36-56. Menlo Park, CA, USA.
- [15] Gruber T., "Toward Principles for design of Ontologies used for Knowledge Sharing", Technical Report KSL-93-04, Knowledge Systems Laboratory, Stanford University, CA, 1993.
- [16] Swartout, William, et al. 1996. "*Toward Distributed Use of Large-Scale Ontologies*". University of Alberta at Calgary. 1996. Disponible en: http://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/Banff_96_final_2.html
- [17] Dorothea Tsatsou, Maria Loli, Vasileios Mezaris (CERTh); Rüdiger Klein, Manuel Kober (FRAUNHOFER); Tomáš Kliegr, Jaroslav Kuchar (UEP); Matei Mancas, Julien Leroy (UMONS); Lyndon Nixon (STI), "Deliverable 4.1 Specification of user profiling and contextualisation, Television Linked to the Web - Integrated Project, LinkedTV, 2012.
- [18] N. F. Noy, C. D. Hafner, "The State of the Art in Ontology Design, A Survey and comparative Review", *AI Magazine*, Fall 1997, pp. 53-74

- [19] Golemati Maria, Katifori Akrivi, Vassilakis Costas, Lepouras Georgios, Halatsis Constantin "Creating an Ontology for the User Profile: Method and Applications"
- [20] M. R. Tazari, M. Grimm, M. Finke, "Modeling User Context", Proceedings of the 10th International Conference on Human-Computer Interaction (HCI2003), Crete, Greece, June 2003
- [21] J. Trajkova, S. Gauch, "Improving Ontology-based User Profiles", Proc. Of RIAO 2004, University of Avignon (Vaucluse), France, April 26-28, 2004, pp. 380-389
- [22] S. Gauch, J. Chaffee, A. Pretschner, "Ontology-Based User Profiles for Search and Browsing, User Modeling and User-Adapted Interaction: The Journal of Personalization Research", Special Issue on User Modeling for Web and Hypermedia Information Retrieval, 2003.
- [23] J. Teevan, S. T. Dumais, and E. Horvitz, "Personalizing Search via automated Analysis of Interests and Activities", Proceedings of SIGIR 2005, ACM Press, August 2005
- [24] B. Kules, User Modeling for Adaptive and Adaptable Software Systems, 2000, Available at <http://www.otal.umd.edu/UUGuide/wmk/>
- [25] A. Kobsa, "User Modelling: Recent work, prospects and hazards", in Adaptive User Interfaces: Principles and Practices, edited by Schneider-Hufschmidt, T. Khme and U. Malinowski, North Holland, 1993.
- [26] J. Fink, A. Kobsa, and A. Nill, "Adaptable and Adaptive Information Access for All Users, Including the Disabled and the Elderly", In User Modeling: Proceedings of the Sixth International Conference, UM97, Vienna, New York, Springer Wien New York, 1997, pp 171 – 173.
- [27] Neches, Robert; Fikes, Richard; Finin, Timothy; Gruber, Thomas; Patil, Ramesh; Senator, Edward & Swartout, William. 1991. "Enabling Technology for Knowledge Sharing". En: Artificial Intelligence Magazine. Vol. 12, num. 3, American Association for Artificial Intelligence Press. 36-56. Menlo Park, CA, USA.
- [28] Gruber T., "Toward Principles for design of Ontologies used for Knowledge Sharing", Technical Report KSL-93-04, Knowledge Systems Laboratory, Stanford University, CA, 1993.

- [29] Swartout, William, et al. 1996. "Toward Distributed Use of Large-Scale Ontologies". University of Alberta at Calgary. 1996. Disponible en: http://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/Banff_96_final_2.html
- [30] D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz, and M. von Wilamowitz-Moellendorff, "Gumo - the general user model ontology." in User Modeling, 2005, pp. 428–432.
- [31] Dorothea Tsatsou, Maria Loli, Vasileios Mezaris (CERTh); Rüdiger Klein, Manuel Kober (FRAUNHOFER); Tomáš Kliegr, Jaroslav Kuchar (UEP); Matei Mancas, Julien Leroy (UMONS); Lyndon Nixon (STI), "Deliverable 4.1 Specification of user profiling and contextualisation, Television Linked to the Web - Integrated Project, LinkedTV, 2012.
- [32] Lenat D.B., Guha R.V., (1990) Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project, Addison-Wesley, Boston.
- [33] Uschold M., King M., (1995) Towards a Methodology for Building Ontologies, in: IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal.
- [34] Gruninger M., Fox M.S., (1995) Methodology for the design and evaluation of ontologies, in: Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal.
- [35] Schreiber Ath., Wielinga B., Jansweijer W.,(1995) The KACTUS view on the 'O' word. Technical Report, ESPRIT Project 8145KACTUS, University of Amsterdam, The Netherlands.
- [36] Gomez-Perez A., M. Fernandez-Lopez, A. de Vicente (1996), Towards a Method to Conceptualize Domain Ontologies, in: ECAI96 Workshop on Ontological Engineering, Budapest, pp. 41–51.
- [37] M. Suarez-Figueroa, NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse, ser. Dissertations in Artificial Intelligence. IOS Press, Incorporated, 2013.
- [38] S. Middleton, N. Shadbolt, and D. Roure, "Ontological user profiling in recommender systems," in ACM Transactions on Information Systems, 2004, pp. 54–88.
- [39] A. Sieg, B. Mobasher, and R. Burke, "Improving the effectiveness of collaborative recommendation with ontology-based user profiles," in Proceedings of the 1st

International Workshop on Information Heterogeneity and Fusion in Recommender Systems, ser. HetRec '10. New York, NY, USA: ACM, 2010, pp. 39–46.

[40] N. P. Andersen, “Reducing cold start problem in the wikipedia recommender system,” Master’s thesis, Technical University of Denmark, DTU Informatics, Asmussens Alle, Building 305, DK-2800 Kgs. Lyngby, Denmark, 2011.

[41] I. Cantador, P. Castells, and A. Bellogín, “An enhanced semantic layer for hybrid recommender systems: Application to news recommendation,” *Int. J. Semant. Web Inf. Syst.*, vol. 7, no. 1, pp. 44–78, Jan. 2011. [Online]. Available: <http://dx.doi.org/10.4018/jswis.2011010103>

[42] V. Saquicela, M. Espinoza, J. Mejia, and B. Villazón-Terrazas, “Reduciendo la sobrecarga de información en usuarios de televisión digital,” in *Proceedings of the Workshop on Semantic Web and Linked Data*, Morelia, México, 2014.

[43] S. Bechhofer, F. van Harmelen, J. Hendler, I. Horrocks, D. L. McGuinness, P. F. Patel-Schneider, and L. A. Stein, “OWL Web Ontology Language Reference,” W3C, Tech. Rep., February 2004.

[44] M. Grüninger and M. S. Fox, “Methodology for the design and evaluation of ontologies,” in *International Joint Conference on Artificial Intelligence (IJCAI95), Workshop on Basic Ontological Issues in Knowledge Sharing*, 1995.

[45] S. Lawrence, “Context in web search,” *IEEE Data Engineering Bulletin*, vol. 23, no. 3, pp. 25–32, 2000.

[46] J. Trajkova and S. Gauch, “Improving ontology-based user profiles,” in *Proc. of RIAO 2004*, 2004, pp. 380–389.

[47] V. Katifori, A. Poggi, M. Scannapieco, T. Catarci, and Y. E. Ioannidis, “Ontopim: how to rely on a personal ontology for personal information management,” in *Semantic Desktop Workshop*, ser. *CEUR Workshop Proceedings*, S. Decker, J. Park, D. Quan, and L. Sauermann, Eds., vol. 175. CEUR-WS.org, 2005.

[48] I. Cantador, P. Castells, and A. Bellogín, “An enhanced semantic layer for hybrid recommender systems: Application to news recommendation,” *Int. J. Semant. Web Inf. Syst.*, vol. 7, no. 1, pp. 44–78, Jan. 2011. [Online]. Available: <http://dx.doi.org/10.4018/jswis.2011010103>

- [49] D. Palmisano and M. Minno, 2010, user and Context Model Specification. Notube Project Deliverable D3.1.
- [50] D. Tsatsou, M. Loli, V. Mezaris, R. Klein, M. Kober, T. Kliegr, J. Kuchar, M. Mancas, J. Leroy, and L. Nixon, 2011, deliverable 4.1 Specification of user profiling and contextualisation. Linked TV Project.
- [51] M. del Carmen Suárez de Figueroa Baonza, "Neon methodology for building ontology networks: Specification, scheduling and reuse," Ph.D. dissertation, Technical University of Madrid, UPM, 2010.
- [52] "Best Practices For Form Design" Luke Wroblewski, 2008
- [53] Architecture Specifications VIRTUOSO SERVER. [Online]. Available: virtuoso.openlinksw.com
- [54] "Enriquecimiento semántico de guías de programación electrónica" Tesis de Grado Universidad de Cuenca. Chica Fernanda, Abad Karina, 2014